

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2005, 10:37:31 ; Search time 11.7951 Seconds
(without alignments)
1180.438 Million cell updates/sec

Title: US-09-762-261C-1_COPY_301_336
Perfect score: 201
Sequence: 1 CSRPNNTKSPMGPRAFYTTGQIIGDIROAHCN 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	201	100.0	866	3	AAY79020 HIV-1 env
2	195	97.0	35	3	AAY79023 HIV-1 RV
3	187	93.0	842	3	AAB69350 HIV-1 non
4	185	92.0	35	2	AAR20883 PND EE378
5	185	92.0	35	2	AAR20953 PND EE669
6	185	92.0	35	2	AAR20853 PND EE322
7	183	91.0	35	2	AAR21134 PND EE228
8	183	91.0	35	2	AAR20911 PND EE533
9	183	91.0	35	2	AAR21135 PND EE228
10	183	91.0	35	2	AAR20912 PND EE533
11	183	91.0	35	2	AAR20910 PND EE533
12	183	91.0	35	2	AAR21136 PND EE228
13	183	91.0	35	2	AAR29113 Group I1
14	183	91.0	45	8	ADR58150 Novel ant
15	182	90.5	39	2	AAR64056 HIV-1 Ba-
16	182	90.5	43	2	AAR63735 HIV-1 JR-
17	182	90.5	45	8	ADR58145 Novel ant
18	182	90.5	60	2	AAR68028 HIV-1 gpl
19	182	90.5	376	6	ABU07702 Viral coa
20	182	90.5	473	4	AAB61502 HIV-1 gpl
21	182	90.5	473	6	ABR57050 HIV-1 JR-
22	182	90.5	473	7	ABR82838 HIV env g
23	182	90.5	473	7	ADF18104 HIV-1 str
24	182	90.5	474	8	ADP74743 HIV-1 iso
25	182	90.5	474	8	ADP74742 HIV-1 iso

26	182	90.5	502	4	AAB82953 HIV-1 JR-
27	182	90.5	506	6	ABU07703 Viral coa
28	182	90.5	506	6	ABU07700 Viral coa
29	182	90.5	507	2	AAR60789 HIV virus
30	182	90.5	507	2	AAR60783 HIV virus
31	182	90.5	556	6	ABU07696 Viral coa
32	182	90.5	579	4	ABU61506 HIV-1 del
33	182	90.5	579	6	ABR57053 HIV-1 JR-
34	182	90.5	579	6	ADF18109 HIV-1 JR-
35	182	90.5	579	8	ADK14401 HIV gpl40
36	182	90.5	590	6	ABU07697 Viral coa
37	182	90.5	643	4	AAB61505 HIV-1 SOS
38	182	90.5	643	6	ABR57052 HIV-1 JR-
39	182	90.5	643	7	ABR82839 HIV env g
40	182	90.5	643	7	ADF18107 HIV-1 JR-
41	182	90.5	643	8	ADK14399 HIV gpl40
42	182	90.5	720	6	ABU07694 Viral coa
43	182	90.5	720	6	ABU07695 Viral coa
44	182	90.5	847	8	ADK14404 HIV mutan
45	182	90.5	847	8	ADK14405 HIV mutan

ALIGNMENTS

RESULT 1
AAY79020
ID AAY79020 standard; protein; 866 AA.
XX
AC AAY79020;
XX
DT 12-SEP-2003 (revised)
DT 05-JUN-2000 (first entry)
XX
DE HIV-1 envelope protein amino acid sequence.
XX
KW HIV-1; envelope protein; env; cross reactive anti-serum; antibody;
KW vaccine; antibody; detect.
XX
OS Human immunodeficiency virus 1.
XX
PN WC200007631-A1.
XX
PD 17-FEB-2000.
XX
PF 04-AUG-1999; 99WO-US017596.
XX
PR 04-AUG-1998; 98US-0095267P.
XX
PA (JACK-) JACKSON FOUND HENRY M.
XX
PI Quinman GV, Zhang PF;
XX
DR WPI; 2000-205578/18.
XX
PT Human immunodeficiency virus-1 envelope protein, useful as a vaccine or
PT immunogenic preparation against HIV-1 infection induces the production of
PT broadly cross-reactive neutralizing anti-serum.
XX
PS Claim 2; Page 33; 54pp; English.
XX
CC This sequence represents a human immunodeficiency virus type 1 (HIV-1)
CC envelope (env) protein amino acid sequence. The invention relates to the
CC HIV-1 env protein or its fragments, which when administered to a mammal,
CC induces the production of broadly cross-reactive neutralising anti-serum
CC against multiple strains of HIV-1. The HIV-1 env protein and its
CC fragments are useful as vaccines, immunogenic compositions or diagnostic
CC reagents. They can also be used to detect anti-HIV-1/HIV-1 neutralizing
CC antibodies in a sample. The amino acid sequence and its fragments or mRNA
CC sequence are useful for generating antibodies in a mammal. In addition, a
CC recombinant delivery vector containing the env amino acid sequence may
CC also be used as a vaccine. (Updated on 12-SEP-2003 to standardise OS
CC field)

```

XX SQ Sequence 866 AA;
Query Match 100.0%; Score 201; DB 3; Length 866;
Best Local Similarity 100.0%; Pred. No. 4.4e-18;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHCN 36
| | | | | | | | | | | | | | | | | | | | | |
Db 301 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHCN 336

RESULT 2
AAY79023
ID AAY79023 standard; protein; 35 AA.
XX AAY79023;
AC
XX
XX 12-SEP-2003 (revised)
DT 05-JUN-2000 (first entry)
XX
XX HIV-1 RV strain envelope protein V3 region peptide sequence #2.
XX
XX HIV-1; envelope protein; env; cross reactive anti-serum; antibody;
KW vaccine; antibody; V3 region; detect.
XX
XX Human immunodeficiency virus 1.
XX
XX WO200007631-A1.
XX
XX 17-FEB-2000.
XX
XX 04-AUG-1999; 99WO-US017596.
XX
XX 04-AUG-1998; 98US-0095267P.
XX
XX (JACK-) JACKSON FOUND HENRY M.
XX
XX Quinnan GV, Zhang PF;
XX
XX WPI; 2000-205578/18.
XX
XX Human immunodeficiency virus-1 envelope protein, useful as a vaccine or
PT immunogenic preparation against HIV-1 infection induces the production of
PT broadly cross-reactive neutralizing anti-serum.
XX
XX Example 5; Page 27; 54pp; English.
XX
XX This sequence represents a human immunodeficiency virus type 1 (HIV-1) R2
CC strain envelope (env) protein V3 region peptide sequence. The invention
CC relates to the HIV-1 env protein or its fragments, which when
CC administered to a mammal, induces the production of broadly cross-
CC reactive neutralising anti-serum against multiple strains of HIV-1. The
CC HIV-1 env protein and its fragments are useful as vaccines, immunogenic
CC compositions or diagnostic reagents. They can also be used to detect anti
CC -HIV-1/HIV-1 neutralizing antibodies in a sample. The amino acid sequence
CC and its fragments or mRNA sequence are useful for generating antibodies
CC in a mammal. In addition, a recombinant delivery vector containing the
CC env amino acid sequence may also be used as a vaccine. (Updated on 12-SEP
CC -2003 to standardise OS field)
XX
XX Sequence 35 AA;
SQ
Query Match 97.0%; Score 195; DB 3; Length 35;
Best Local Similarity 100.0%; Pred. No. 1e-18;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHC 35
| | | | | | | | | | | | | | | | | | | | | |
Db 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHC 35

RESULT 3
AAB69350
ID AAB69350 standard; protein; 842 AA.
XX AAB69350;
AC
XX
XX 12-SEP-2003 (revised)
DT 20-APR-2001 (first entry)
XX
XX HIV-1 non-subtype B clone 93BR029-4 env protein.
XX
XX HIV-1; human immunodeficiency virus; non-subtype B; gag; env; vpu;
KW vif; vpr; tat; rev; nef; vaccine.
XX
XX Human immunodeficiency virus 1.
XX
XX WO200026416-A1.
XX
XX 11-MAY-2000.
XX
XX 25-OCT-1999; 99WO-US024837.
XX
XX 02-NOV-1998; 98US-00184418.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Hahn BH, Shaw GM, Gao F;
XX
XX WPI; 2000-365651/31.
XX
XX Novel genomic nucleic acids of non-subtype B human immunodeficiency virus
PT type 1 useful for detecting and treating AIDS comprises a specific
PT nucleotide sequence.
XX
XX Claim 41; Fig 21; 131pp; English.
XX
XX The present invention provides the protein and coding sequences for a
CC number of human immunodeficiency virus (HIV) type 1 non-subtype B
CC isolates. The sequences shown include the near full-length coding
CC sequences from each isolate, and the env, pol, vif, vpr, vpu, gag, tat,
CC rev and nef proteins. These can be used to detect the presence of HIV-1
CC in a sample and to produce antibodies against non-subtype B HIV-1 virus.
CC These antibodies can be used in vaccines to prevent and treat HIV
CC infection. (Updated on 12-SEP-2003 to standardise OS field)
XX
XX Sequence 842 AA;
SQ
Query Match 93.0%; Score 187; DB 3; Length 842;
Best Local Similarity 88.9%; Pred. No. 3.2e-16;
Matches 32; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHCN 36
| | | | | | | | | | | | | | | | | | | | | |
Db 298 CTRPNNTKRSIPGPGRAFYTTGQIIGDIRQAHCN 333
| | | | | | | | | | | | | | | | | | | | | |

RESULT 4
AAR20883
ID AAR20883 standard; protein; 35 AA.
XX AAR20883;
AC
XX
XX 01-JUN-1992 (first entry)
XX
XX PND EE378-3.
XX
XX HIV; PND; principal neutralising determinant; Omp; Neisseria;
KW outer membrane proteosome; AIDS; vaccine; envelope gene.
XX
XX Human immunodeficiency virus.
XX
XX EP471407-A.
XX
XX 19-FEB-1992.

```

```

XX 07-AUG-1991; 91EP-00202025.
XX 13-AUG-1990; 90US-00566638.
XX 13-AUG-1990; 90US-00566654.
XX 13-AUG-1990; 90US-00566656.
XX (MERI ) MERCK & CO INC.
XX Lewis JA, Davide JP, Waterbury JA;
XX WPI: 1992-058471/08.
XX N-PSDB; AAQ21298.
XX New antigenic conjugate of HIV major neutralisation determinant -
XX covalently linked to outer membrane proteosome of Neisseria, useful as
XX vaccine and in treating and preventing HIV infections, AIDS and ARC.
XX Claim 1; Page 175 + 117-118; 177pp; English.
XX The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78
XX encode PNDs of HIV (envelope protein) which are useful as immunogens for
XX AIDS vaccines, partic. in the form of conjugates. To prepare antigenic
XX conjugates, PND (of HIV) and Omp (of Neisseria) are prepd. separately,
XX then linked by cross-linking agents, monogeneric spacers or bigeneric
XX spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids
XX or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,
XX Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or
XX post- exposure to prevent or treat HIV infection or disease, and are
XX capable of producing an immune response specific for the immunogen. See
XX also AAQ21052-61, AAQ21194-311 and AAQ21321-78
XX Sequence 35 AA;
XX Query Match 92.0%; Score 185; DB 2; Length 35;
XX Best Local Similarity 91.4%; Pred. No. 2.3e-17; Indels 0; Gaps 0;
XX Matches 32; Conservative 3; Mismatches 0;
XX QY 1 CSRPNNTKRSIPMGPGRAFYTGTGQIIGDIRQAHC 35
XX |:|||||:|||||:|||||:|||||:|||||:
XX 1 CTRPNNTKRSIPGPGRAFYTGTGQIIGDIRQAHC 35
XX Ddb
XX RESULT 5
XX IID AAR20953
XX XX AAR20953 standard; protein; 35 AA.
XX XX AAR20953;
XX AC
XX 01-JUN-1992 (first entry)
XX DE PND EE669-3.
XX XX HIV; PND; principal neutralising determinant; Omp; Neisseria;
XX XX outer membrane proteosome; AIDS; vaccine; envelope gene.
XX XX Human immunodeficiency virus.
XX XX EP471407-A.
XX XX 19-FEB-1992.
XX XX 07-AUG-1991; 91EP-00202025.
XX XX 13-AUG-1990; 90US-00566638.
XX XX 13-AUG-1990; 90US-00566654.
XX XX 13-AUG-1990; 90US-00566656.
XX XX (MERI ) MERCK & CO INC.
XX XX Lewis JA, Davide JP, Waterbury JA;
XX XX WPI: 1992-058471/08.
XX XX
XX

```

CC then linked by cross-linking agents, monogeneric spacers or bigeneric
 CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids
 CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,
 CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or
 CC post- exposure to prevent or treat HIV infection or disease, and are
 CC capable of producing an immune response specific for the immunogen. See
 CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78

XX SQ Sequence 35 AA;

Query Match 92.0%; Score 185; DB 2; Length 35;
 Best Local Similarity 91.4%; Pred. No. 2.3e-17;
 Matches 32; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHC 35
 Db 1 CTRPNNTKRSIPGPGRAFYTTGQIIGDIRQAHC 35

RESULT 7

AAR21134
 ID AAR21134 standard; protein; 35 AA.

XX AAR21134;

XX 01-JUN-1992 (first entry)

XX PND EE228-1.

XX HIV; PND; principal neutralising determinant; Omp; Neisseria;
 KW outer membrane proteosome; AIDS; vaccine; envelope gene.

XX Human immunodeficiency virus.

XX EP471407-A.

XX 19-FEB-1992.

XX 07-AUG-1991; 91EP-00202025.

XX 13-AUG-1990; 90US-00566638.

XX 13-AUG-1990; 90US-00566654.

XX 13-AUG-1990; 90US-00566656.

XX (MERI) MERCK & CO INC.

XX Lewis JA, Davide JP, Waterbury JA;

XX WPI; 1992-058471/08.

XX N-PSDB; AAQ21232.

XX New antigenic conjugate of HIV major neutralisation determinant -

XX covalently linked to outer membrane proteosome of Neisseria, useful as

XX vaccine and in treating and preventing HIV infections, AIDS and ARC.

XX Claim 1; Page 175 + 64; 177pp; English.

XX The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78
 CC encode PNDs of HIV (envelope protein) which are useful as immunogens for
 CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic
 CC conjugates, PND (of HIV) and Omp (of Neisseria) are prep'd. separately,
 CC then linked by cross-linking agents, monogeneric spacers or bigeneric
 CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids
 CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,
 CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or
 CC post- exposure to prevent or treat HIV infection or disease, and are
 CC capable of producing an immune response specific for the immunogen. See
 CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78

XX SQ Sequence 35 AA;

Query Match 91.0%; Score 183; DB 2; Length 35;
 Best Local Similarity 91.4%; Pred. No. 4.2e-17;

Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHC 35
 Db 1 CTRPNNTKRSIPGPGRAFYTTGQIIGDIRQAHC 35

RESULT 8

AAR20911
 ID AAR20911 standard; protein; 35 AA.

XX AAR20911;

XX 01-JUN-1992 (first entry)

XX PND EE533-2.

XX HIV; PND; principal neutralising determinant; Omp; Neisseria;
 KW outer membrane proteosome; AIDS; vaccine; envelope gene.

XX Human immunodeficiency virus.

XX EP471407-A.

XX 19-FEB-1992.

XX 07-AUG-1991; 91EP-00202025.

XX 13-AUG-1990; 90US-00566638.

XX 13-AUG-1990; 90US-00566654.

XX 13-AUG-1990; 90US-00566656.

XX (MERI) MERCK & CO INC.

XX Lewis JA, Davide JP, Waterbury JA;

XX WPI; 1992-058471/08.

XX N-PSDB; AAQ21335.

XX New antigenic conjugate of HIV major neutralisation determinant -

XX covalently linked to outer membrane proteosome of Neisseria, useful as

XX vaccine and in treating and preventing HIV infections, AIDS and ARC.

XX Claim 1; Page 175 + 140; 177pp; English.

XX The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78
 CC encode PNDs of HIV (envelope protein) which are useful as immunogens for
 CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic
 CC conjugates, PND (of HIV) and Omp (of Neisseria) are prep'd. separately,
 CC then linked by cross-linking agents, monogeneric spacers or bigeneric
 CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids
 CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,
 CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or
 CC post- exposure to prevent or treat HIV infection or disease, and are
 CC capable of producing an immune response specific for the immunogen. See
 CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78

XX SQ Sequence 35 AA;

Query Match 91.0%; Score 183; DB 2; Length 35;
 Best Local Similarity 91.4%; Pred. No. 4.2e-17;
 Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHC 35

Db 1 CTRPNNTKRSIPGPGRAFYTTGQIIGDIRQAHC 35

RESULT 9

AAR21135
 ID AAR21135 standard; protein; 35 AA.

XX AAR21135;

DR WPI; 1992-058471/08.
 DR N-PSDB; AAQ21334.
 XX
 PT New antigenic conjugate of HIV major neutralisation determinant -
 PT covalently linked to outer membrane proteosome of Neisseria, useful as
 PT vaccine and in treating and preventing HIV infections, AIDS and ARC.
 XX
 PS Claim 1; Page 175 + 139-140; 177pp; English.
 XX
 CC The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78
 CC encode PNDs of HIV (envelope protein) which are useful as immunogens for
 CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic
 CC conjugates, PND (of HIV) and Omp (of Neisseria) are prep. separately,
 CC then linked by cross-linking agents, monogeneric spacers or bigeneric
 CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids
 CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,
 CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or
 CC post- exposure to prevent or treat HIV infection or disease, and are
 CC capable of producing an immune response specific for the immunogen. See
 CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78
 XX
 SQ Sequence 35 AA;
 CC
 CC Query Match 91.0%; Score 183; DB 2; Length 35;
 CC Best Local Similarity 91.4%; Pred. No. 4.2e-17;
 CC Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CSRPNNTKRSIPMGCGRAFYTTGQIIIGDIRQAHG 35
 DB 1 CTRPNNTKRSIPMGCGRAFYTTGQIIIGDIRQAHG 35
 CC
 CC RESULT 12
 CC AAR21136
 ID AAR21136 standard; protein; 35 AA.
 AC AAR21136;
 XX
 XX 01-JUN-1992 (first entry)
 DT
 XX PND BR228-3.
 DE
 XX HIV; PND; principal neutralising determinant; Omp; Neisseria;
 KW outer membrane proteosome; AIDS; vaccine; envelope gene.
 KW
 XX Human immunodeficiency virus.
 OS
 XX EP471407-A.
 PN
 XX 19-FEB-1992.
 PD
 XX 07-AUG-1991; 91EP-00202025.
 PF
 XX 13-AUG-1990; 90US-00566638.
 PR
 XX 13-AUG-1990; 90US-00566654.
 PR
 XX 13-AUG-1990; 90US-00566656.
 XX
 XX (MERI) MERCK & CO INC.
 PA
 XX Lewis JA, Davide JP, Waterbury JA;
 PI
 XX WPI; 1992-058471/08.
 DR
 DR N-PSDB; AAQ21234.
 XX
 XX New antigenic conjugate of HIV major neutralisation determinant -
 PT covalently linked to outer membrane proteosome of Neisseria, useful as
 PT vaccine and in treating and preventing HIV infections, AIDS and ARC.
 XX
 PS Claim 1; Page 175 + 66; 177pp; English.
 XX
 CC The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78
 CC encode PNDs of HIV (envelope protein) which are useful as immunogens for
 CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic

CC conjugates, PND (of HIV) and Omp (of Neisseria) are prep. separately,
 CC then linked by cross-linking agents, monogeneric spacers or bigeneric
 CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids
 CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,
 CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or
 CC post- exposure to prevent or treat HIV infection or disease, and are
 CC capable of producing an immune response specific for the immunogen. See
 CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78
 XX
 SQ Sequence 35 AA;
 CC
 CC Query Match 91.0%; Score 183; DB 2; Length 35;
 CC Best Local Similarity 91.4%; Pred. No. 4.2e-17;
 CC Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CSRPNNTKRSIPMGCGRAFYTTGQIIIGDIRQAHG 35
 DB 1 CTRPNNTKRSIPMGCGRAFYTTGQIIIGDIRQAHG 35
 CC
 CC RESULT 13
 CC AAR29113
 ID AAR29113 standard; protein; 35 AA.
 AC AAR29113;
 XX
 XX 25-MAR-2003 (revised)
 DT
 XX 02-APR-1993 (first entry)
 DT
 XX Group II HIV gp120 PND peptide 8926C.
 DE
 XX
 KW Principal Neutralising determinant; Human immunodeficiency virus;
 KW vaccine; Robson's analytical method; polymerase chain reaction;
 KW Garnier-Osguthorpe-Robson method; GOR method; secondary structure.
 KW
 XX Human immunodeficiency virus.
 OS
 XX
 FH Key Location/Qualifiers
 FT Region 8..14
 FT /note= "classification of PND peptides is determined by
 FT the predicted secondary structure of this region - see
 FT comments"
 FT 15..18
 FT /note= "conserved PND motif"
 FT
 XX EP516135-A2.
 PN
 XX 02-DEC-1992.
 PD
 XX 29-MAY-1992; 92EP-00109072.
 PF
 XX 31-MAY-1991; 91JP-00129224.
 PR
 XX (KAGA) CHEMO SERO THERAPEUTIC RES INS.
 PA
 XX Eda Y, Shiosaki K, Osatomi K, Tokiyoshi S;
 PI
 XX WPI; 1992-400517/49.
 DR
 XX Principle neutralising determinant peptide(s) of HIV gp120 protein - used
 PT for diagnosing, preventing and treating HIV infection.
 PT
 XX Example 1; Page 10; 26pp; English.
 PS
 XX DNA encoding HIV PND peptides was PCR amplified using genomic DNA from
 CC HIV-infected peripheral blood mononuclear cells as template. The
 CC amplified fragments were fused to beta-galactosidase coding sequence.
 CC E.coli transformants were cultured to produce the fusion protein. The
 CC expressed PND proteins were divided into groups based on their reactivity
 CC with neutralising antibodies and their amino acid sequence. The amino
 CC acid sequence was analysed using Robson's analytical program for protein
 CC secondary structure. Five groups were identified and 90% of all
 CC previously reported PND peptides were included in 3 main groups (i.e.

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2005, 10:57:57 ; Search time 2.72785 Seconds
(without alignments)
1269.790 Million cell updates/sec

Title: US-09-762-261C-1_COPY_301_336
Perfect score: 201
Sequence: 1 CSRPNNTKRSIPMGPGAFYTTGGIIGDIRQAHCN 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	182	90.5	77	2 S35862	env polyprotein -
2	182	90.5	110	2 A46410	env polyprotein -
3	182	90.5	847	2 T09448	env polyprotein -
4	182	90.5	847	2 S13289	env polyprotein -
5	181	90.0	76	2 S35859	env polyprotein -
6	181	90.0	77	2 S35820	env polyprotein -
7	181	90.0	77	2 S35819	env polyprotein -
8	180	89.6	35	2 PC2296	V3 domain peptide
9	180	89.6	443	2 C41621	env polyprotein P
10	179	89.1	77	2 S35834	env polyprotein -
11	179	89.1	77	2 S35836	env polyprotein -
12	179	89.1	77	2 S35835	env polyprotein -
13	178	88.6	77	2 S35848	env polyprotein -
14	178	88.6	77	2 S35849	env polyprotein -
15	178	88.6	77	2 S35857	env polyprotein -
16	177	88.1	46	2 G01847	env polyprotein -
17	177	88.1	77	2 S35823	env polyprotein -
18	177	88.1	77	2 S35809	env polyprotein -
19	177	88.1	110	2 B46410	env polyprotein -
20	177	88.1	110	2 C46410	env polyprotein -
21	177	88.1	290	2 S25940	env polyprotein -
22	177	88.1	445	2 A41621	env polyprotein M
23	176	87.6	35	2 PC2295	V3 domain peptide
24	176	87.6	77	2 S35806	env polyprotein -
25	176	87.6	77	2 S35841	env polyprotein -
26	176	87.6	77	2 S35816	env polyprotein -
27	176	87.6	77	2 S35804	env polyprotein -
28	176	87.6	77	2 S35850	env polyprotein -
29	176	87.6	843	1 H44001	env polyprotein pr

30	175	87.1	71	2 S42918	env polyprotein -
31	175	87.1	77	2 S35821	env polyprotein -
32	174	86.6	77	2 S35824	env polyprotein -
33	174	86.6	77	2 S35825	env polyprotein -
34	174	86.6	110	2 D46410	env polyprotein -
35	174	86.6	852	1 VCLJBR	env polyprotein -
36	173	86.1	71	2 S42904	env polyprotein -
37	173	86.1	71	2 S42905	env polyprotein -
38	173	86.1	71	2 S42898	env polyprotein -
39	173	86.1	77	2 S35815	env polyprotein -
40	172	85.6	506	2 A40218	env polyprotein -
41	171	85.1	35	2 PC2297	V3 domain peptide
42	171	85.1	77	2 S35800	env polyprotein -
43	171	85.1	77	2 S35801	env polyprotein -
44	171	85.1	299	2 S60529	env polyprotein -
45	171	85.1	855	1 VCLJAJ	env polyprotein pr

ALIGNMENTS

RESULT 1

S35862
env polyprotein - human immunodeficiency virus type 1 (strain CSF2951) (fragment)
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
C:Accession: S35862; S35860; S35863

R;Chiodi, F.
submitted to the EMBL Data Library, June 1993

A:Reference number: S35800
A:Accession: S35862
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-77 <CH1>
A:Cross-references: UNIPROT:Q77544; UNIPROT:Q77542; UNIPROT:Q77545; EMBL:Z23246; NID:g311;
A:Experimental source: strain CSF2951 clone3; strain CSF2951 clone4
A:Accession: S35860
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-77 <CH2>
A:Cross-references: EMBL:Z23244; NID:g313631; PIDN:CAA80762.1; PID:g313632; EMBL:Z23247;
A:Experimental source: strain CSF2951 clone11; strain CSF2951 clone4
C:Superfamily: type E retrovirus env polyprotein

Query Match 90.5%; Score 182; DB 2; Length 77;
Best Local Similarity 88.9%; Pred. No. 2.8e-18;
Matches 32; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CSRPNNTKRSIPMGPGAFYTTGGIIGDIRQAHCN 36
|:|||||:|||||:|||||:|||||:|||||
Db 26 CTRPNNTKRSINIGPGAFYTTGGIIGDIRQAHCN 61

RESULT 2

A46410
env polyprotein (V3 domain) - human immunodeficiency virus type 1 (fragment)
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A46410
R;McNairney, T.; Hornickova, Z.; Markham, R.; Birdwell, A.; Arens, M.; Saah, A.; Ratner, Proc. Natl. Acad. Sci. U.S.A. 89, 10247-10251, 1992
A:Title: Relationship of human immunodeficiency virus type 1 sequence heterogeneity to st
A:Reference number: A46410; MUID:93066216; PMID:1438212
A:Accession: A46410
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-110 <MCN>
A:Cross-references: UNIPROT:Q9PY06
A:Experimental source: subject S1
A:Note: sequence extracted from NCBI backbone (NCBIP:117724)
C:Superfamily: type E retrovirus env polyprotein

Query Match 90.5%; Score 182; DB 2; Length 110;

A;Accession: PC2296
A;Molecule type: protein
A;Residues: 1-35 <SHE>
A;Cross-references: UNIPROT:Q70826; UNIPROT:Q70831
C;Superfamily: type E retrovirus env polyprotein

Query Match 89.6%; Score 180; DB 2; Length 35;
Best Local Similarity 91.4%; Pred. No. 2.3e-18;
Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIHMGPGRAFYTTGQIIGDIRQAHN 35
DB 1 CTRPNNTKSIHMGPGRAFYTTGQIIGDIRQAHN 35

RESULT 9

env polyprotein P - human immunodeficiency virus type 1 (fragment)
N;Alternate names: coat polyprotein
N;Contains: amino end of coat protein gp41; carboxyl end of coat protein gp120
C;Species: human immunodeficiency virus type 1, HIV-1
A;Note: host Homo sapiens (man)
C;Date: 11-Feb-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: C41621
R;Burger, H.; Weiser, B.; Flaherty, K.; Gullia, J.; Nguyen, P.N.; Gibbs, R.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 11236-11240, 1991
A;Title: Evolution of human immunodeficiency virus type 1 nucleotide sequence diversity
A;Reference number: A41621; MUID:92107924; PMID:1763038
A;Accession: C41621
A;Molecule type: DNA
A;Residues: 1-443 <BUR>
A;Cross-references: UNIPROT:Q80023; GB:M77230; NID:G328631; PIDN:AA03792.1; PID:G555015
A;Note: this virus was isolated from the mother's sexual partner
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot
F;1-251/Product: coat protein gp120 (fragment) #status predicted <GP1>
F;252-443/Product: coat protein gp41 (fragment) #status predicted <GP2>
F;424-443/Domain: transmembrane #status predicted <TM>
F;9,23,36,48,78,101,107,131,137,143,147,153,188,200,203,351,356,365,377/Binding site: ca

Query Match 89.6%; Score 180; DB 2; Length 443;
Best Local Similarity 88.9%; Pred. No. 3.3e-17;
Matches 32; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIHMGPGRAFYTTGQIIGDIRQAHN 36
DB 43 CIRPNNTKSIHMGPGRAFYTTGQIIGDIRQAHN 78

RESULT 10

S35834
envelope protein - human immunodeficiency virus type 1 (fragment)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
C;Accession: S35834; S35837
R;Chiodi, F.

submitted to the EMBL Data Library, June 1993

A;Reference number: S35800
A;Accession: S35834
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-77 <CHI>
A;Cross-references: UNIPROT:Q77519; UNIPROT:Q77516; EMBL:Z23216; EMBL:Z23219
C;Superfamily: type E retrovirus env polyprotein

Query Match 89.1%; Score 179; DB 2; Length 77;
Best Local Similarity 88.9%; Pred. No. 7.4e-18;
Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIHMGPGRAFYTTGQIIGDIRQAHN 36
|:|||||||||:|||||||||:|||||||||:|||||||||

DB 26 CTRPNNTKSIHMGPGRAFYATGDIIGDIRQAHN 61

RESULT 11

S35836
envelope protein - human immunodeficiency virus type 1 (fragment)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
C;Accession: S35836
R;Chiodi, F.

submitted to the EMBL Data Library, June 1993

A;Reference number: S35800
A;Accession: S35836
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-77 <CHI>
A;Cross-references: UNIPROT:Q77518; EMBL:Z23218
C;Superfamily: type E retrovirus env polyprotein

Query Match 89.1%; Score 179; DB 2; Length 77;
Best Local Similarity 88.9%; Pred. No. 7.4e-18;
Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIHMGPGRAFYTTGQIIGDIRQAHN 36
DB 26 CTRPNNTKSIHMGPGRAFYATGDIIGDIRQAHN 61

RESULT 12

S35835
envelope protein - human immunodeficiency virus type 1 (fragment)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
C;Accession: S35835
R;Chiodi, F.

submitted to the EMBL Data Library, June 1993

A;Reference number: S35800
A;Accession: S35835
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-77 <CHI>
A;Cross-references: UNIPROT:Q77517; EMBL:Z23217
C;Superfamily: type E retrovirus env polyprotein

Query Match 89.1%; Score 179; DB 2; Length 77;
Best Local Similarity 88.9%; Pred. No. 7.4e-18;
Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIHMGPGRAFYTTGQIIGDIRQAHN 36
DB 26 CTRPNNTKSIHMGPGRAFYATGDIIGDIRQAHN 61

RESULT 13

S35848
envelope protein - human immunodeficiency virus type 1 (fragment)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
C;Accession: S35848
R;Chiodi, F.

submitted to the EMBL Data Library, June 1993

A;Reference number: S35800
A;Accession: S35848
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-77 <CHI>
A;Cross-references: UNIPROT:Q77530; EMBL:Z23232; NID:G313607; PIDN:CAA0750.1; PID:G31361
C;Superfamily: type E retrovirus env polyprotein

Query Match 88.6%; Score 178; DB 2; Length 77;
Best Local Similarity 86.1%; Pred. No. 1e-17;
Matches 31; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTCGIIIGDIRQAHCN 36
|:||||| |:||||| |:||||| |:|||||
Db 26 CTRPNNTRKGIHIGPGRAFYTTCGIIIGDIRQAHCN 61

RESULT 14

S35849
envelope protein - human immunodeficiency virus type 1 (fragment)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
C;Accession: S35849
R;Chiodi, F.
submitted to the EMBL Data Library, June 1993
A;Reference number: S35800
A;Accession: S35849
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-77 <CHI>
A;Cross-references: UNIPROT:Q77531; EMBL:Z23233; NID:g313609; PIDN:CAA80751.1; PID:g3136
C;Superfamily: type E retrovirus env polyprotein

Query Match 88.6%; Score 178; DB 2; Length 77;
Best Local Similarity 86.1%; Pred. No. 1e-17;
Matches 31; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTCGIIIGDIRQAHCN 36
|:||||| |:||||| |:||||| |:|||||
Db 26 CTRPNNTRKGIHIGPGRAFYTTCGIIIGDIRQAHCN 61

RESULT 15

S35857
envelope protein - human immunodeficiency virus type 1 (fragment)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
C;Accession: S35857
R;Chiodi, F.
submitted to the EMBL Data Library, June 1993
A;Reference number: S35800
A;Accession: S35857
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-77 <CHI>
A;Cross-references: UNIPROT:Q77539; EMBL:Z23241; NID:g313625; PIDN:CAA80759.1; PID:g3136
C;Superfamily: type E retrovirus env polyprotein

Query Match 88.6%; Score 178; DB 2; Length 77;
Best Local Similarity 86.1%; Pred. No. 1e-17;
Matches 31; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTCGIIIGDIRQAHCN 36
|:||||| |:||||| |:||||| |:|||||
Db 26 CTRPNNTRKSWIGPGRAFYTTCGIIIGDIRQAHCN 61

Search completed: April 11, 2005, 11:15:28
Job time : 2.72785 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2005, 10:39:11 ; Search time 11.5261 Seconds
(without alignments)
1599.394 Million cell updates/sec

Title: US-09-762-261C-1_COPY_301_336

Perfect score: 201
Sequence: 1 CSRPNNTKSIIPMGPGRAFYTTGQIIGDIRQAHCN 36

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	201	100.0	866	2 Q9WPZ4	Q9WPZ4 human immun
2	198	98.5	91	2 Q39362	Q39362 human immun
3	198	98.5	91	2 Q69692	Q69692 human immun
4	195	97.0	85	2 Q90956	Q90956 human immun
5	195	97.0	88	2 Q9DXK1	Q9DXK1 human immun
6	195	97.0	171	2 Q6UA59	Q6UA59 human immun
7	194	96.5	92	2 Q90205	Q90205 human immun
8	194	96.5	92	2 Q90206	Q90206 human immun
9	194	96.5	92	2 Q79067	Q79067 human immun
10	194	96.5	112	2 Q9DPV9	Q9DPV9 human immun
11	194	96.5	112	2 Q9DPW0	Q9DPW0 human immun
12	194	96.5	112	2 Q9DPW1	Q9DPW1 human immun
13	194	96.5	112	2 Q9DPW2	Q9DPW2 human immun
14	194	96.5	112	2 Q9DPW3	Q9DPW3 human immun
15	194	96.5	113	2 Q71020	Q71020 human immun
16	194	96.5	229	2 Q40047	Q40047 human immun
17	194	96.5	229	2 Q40049	Q40049 human immun
18	194	96.5	229	2 Q40051	Q40051 human immun
19	194	96.5	230	2 Q40055	Q40055 human immun
20	194	96.5	235	2 Q40050	Q40050 human immun
21	194	96.5	235	2 Q40052	Q40052 human immun
22	193	96.0	80	2 Q9Q677	Q9Q677 human immun
23	193	96.0	119	2 Q75417	Q75417 human immun
24	193	96.0	119	2 Q75422	Q75422 human immun
25	193	96.0	119	2 Q75426	Q75426 human immun
26	193	96.0	120	2 Q75415	Q75415 human immun
27	193	96.0	120	2 Q75421	Q75421 human immun
28	193	96.0	170	2 Q8AEX6	Q8AEX6 human immun
29	193	96.0	311	2 Q72513	Q72513 human immun
30	192	95.5	35	2 Q78198	Q78198 human immun
31	192	95.5	74	2 Q8A1V5	Q8A1V5 human immun

RESULT 1
Q9WPZ4 ID Q9WPZ4 PRELIMINARY; PRT; 866 AA.
AC Q9WPZ4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Envelope protein.
DE Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
RN [1]
RP MEDLINE=99236722; PubMed=10221533; DOI=10.1089/08922299311088;
RA Quinlan G.V. Jr., Zhang P.F., Fu D.W., Dong M., Alter H.J.;
RT "Expression and characterization of HIV type 1 envelope protein associated with a broadly reactive neutralizing antibody response.";
RL AIDS Res. Hum. Retroviruses 15:561-570(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Quinlan G.V. Jr., Zhang P.F., Fu D.W., Dong M., Alter H.J.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF128126; AD40637.3; -.
DR HSSP; P04578; 1DLB.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 866 AA; 98113 MW; 6E76021833P2EACD CRC64;

Query Match 100.0%; Score 201; DB 2; Length 866;
Best Local Similarity 100.0%; Pred. No. 1e-19;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIIPMGPGRAFYTTGQIIGDIRQAHCN 36
|||||
Db 301 CSRPNNTKSIIPMGPGRAFYTTGQIIGDIRQAHCN 336

RESULT 2
Q39362 ID Q39362 PRELIMINARY; PRT; 91 AA.
AC Q39362;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).

32 192 95.5 74 2 Q8A1W7 Q8A1W7 human immun
33 192 95.5 111 2 Q9Q865 Q9Q865 human immun
34 192 95.5 198 2 Q6SY25 Q6SY25 human immun
35 192 95.5 198 2 Q6SY26 Q6SY26 human immun
36 192 95.5 198 2 Q6SY27 Q6SY27 human immun
37 192 95.5 198 2 Q6SY28 Q6SY28 human immun
38 192 95.5 198 2 Q6SY29 Q6SY29 human immun
39 191 95.0 53 2 Q9PXN8 Q9PXN8 human immun
40 191 95.0 54 2 Q76143 Q76143 human immun
41 191 95.0 68 2 Q9QJ11 Q9QJ11 human immun
42 191 95.0 73 2 Q92991 Q92991 human immun
43 191 95.0 73 2 Q92992 Q92992 human immun
44 191 95.0 75 2 Q90U52 Q90U52 human immun
45 191 95.0 75 2 Q90U56 Q90U56 human immun

ALIGNMENTS

```

GN Name-env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;

[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98242906; PubMed=9583604;
RA Lukashov V.V., Op de Coul E.L., Coutinho R.A., Goudsmit J.;
RT "HIV-1 strains specific for Dutch injecting drug users in
heterosexually infected individuals in The Netherlands.";
RL AIDS 12:635-641(1998).
DR EMBL; AF032157; AAC59382.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 91
FT NON_TER 91 91
SQ SEQUENCE 91 AA; 10363 MW; 0D9690FA2767592E CRC64;

Query Match 98.5%; Score 198; DB 2; Length 91;
Best Local Similarity 97.2%; Pred. No. 2.4e-20;
Matches 35; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTGTGIIIGDIRQAHCN 36
Db 27 CTRPNNTKRSIPMGPGRAFYTGTGIIIGDIRQAHCN 62

RESULT 3
Q69692 ID Q69692 PRELIMINARY; PRT; 91 AA.
AC Q69692;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein gp120 (Fragment).
GN Name-env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
[1]
RP SEQUENCE FROM N.A.
RX STRAIN=R923572;
RA Lukashov V.V., Kuiken C.L., Boer K., Goudsmit J.;
RT "HIV type 1 subtypes in The Netherlands circulating among women
originating from AIDS-endemic regions.";
RL AIDS Res. Hum. Retroviruses 12:951-953(1996).
DR EMBL; L76886; AAC37946.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 91
FT NON_TER 91 91
SQ SEQUENCE 91 AA; 10333 MW; 17273BE096D7593E CRC64;

Query Match 98.5%; Score 198; DB 2; Length 91;
Best Local Similarity 97.2%; Pred. No. 2.4e-20;
Matches 35; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTGTGIIIGDIRQAHCN 36
Db 27 CTRPNNTKRSIPMGPGRAFYTGTGIIIGDIRQAHCN 62

```

```

RESULT 4
O90956 ID O90956 PRELIMINARY; PRT; 85 AA.
AC O90956;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp120 (Fragment).
GN Name-env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
[1]
RP SEQUENCE FROM N.A.
RA Heyndrickx L., Janssens W., Coppens S., Vereecken K., Willems B.,
Fransen K., Colebunders R., Vandenbrueene M., Van der Groen G.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ228189; CAA12819.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Glycoprotein; Transmembrane.
FT NON_TER 1 85
FT NON_TER 85 85
SQ SEQUENCE 85 AA; 9889 MW; 01BB36F634EA0AFB CRC64;

Query Match 97.0%; Score 195; DB 2; Length 85;
Best Local Similarity 94.4%; Pred. No. 5.9e-20;
Matches 34; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTGTGIIIGDIRQAHCN 36
Db 29 CTRPNNTKRSIPMGPGRAFYTGTGIIIGDIRQAHCN 64

RESULT 5
Q9DXK1 ID Q9DXK1 PRELIMINARY; PRT; 88 AA.
AC Q9DXK1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN Name-env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21084776; PubMed=11216936;
RA Op de Coul E.L.M., Prins M., Cornelissen M., van der Schoot A.,
Boufassa F., Brettler R.P., Hernandez-Aguado L., Schiffer V.,
McMenamin J., Rezza G., Robertson R., Zangerle R., Goudsmit J.,
Coutinho R.A., Lukashov V.V.;
RT "Using phylogenetic analysis to trace HIV-1 migration among western
European injecting drug users seroconverting from 1984 to 1997.";
RL AIDS 15:257-266(2001).
DR EMBL; AF307295; AAG32463.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 88
FT NON_TER 88 88
SQ SEQUENCE 88 AA; 9925 MW; DD523F6D428DB9E4 CRC64;

Query Match 97.0%; Score 195; DB 2; Length 88;

```

KW	AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.							
FT	NON TER 1							
FT	NON TER 92							
SQ	SEQUENCE 92 AA; 10466 MW; 3395C4528B9A2E51 CRC64;							
	Query Match 96.5%; Score 194; DB 2; Length 92;							
	Best Local Similarity 94.4%; Pred. No. 9e-20; Indels 0; Gaps 0;							
	Matches 34; Conservative 2; Mismatches 0; Indels 0; Gaps 0;							
QY	1 CSRPNNTKRSIPMGPGRAFYYTGGIIGDIRQAHCN 36 : : : : :							
Db	27 CTRPNNTRKSIIPIGPGRAFYYTGGIIGDIRQAHCN 62 : : : : :							
RESULT 8								
C90206	PRELIMINARY; PRT; 92 AA.							
ID	O90206 PRELIMINARY; PRT; 92 AA.							
AC	O90206; (TrEMBLrel. 08, Created)							
DT	01-NOV-1998 (TrEMBLrel. 08, Last sequence update)							
DT	01-NOV-1998 (TrEMBLrel. 08, Last sequence update)							
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)							
DE	Envelope glycoprotein (Fragment).							
GN	Name=env;							
OS	Human immunodeficiency virus 1.							
OC	Viruses; Retrovirdae; Retroviridae; Lentivirus.							
OX	NCBI_TaxID=11676;							
RN	[1]							
RP	SEQUENCE FROM N.A.							
RX	MEDLINE=99039932; PubMed=9824326;							
RA	Goudamit J., Lukashov V.V., van Ameijden E.J., Zorgdrager F., van den Burg R., Cornelissen M.; "Impact of sexual versus parenteral transmission events on the evolution of the gag and env genes of HIV type 1."; AIDS Res. Hum. Retroviruses 14:1483-1486(1998).							
DR	EMBL: AF071286; AAC71738.1; -. GO: GO:0016021; C:integral to membrane; IEA. DR	GO: GO:0019028; C:viral capsid; IEA. DR	GO: GO:0019031; C:viral envelope; IEA. DR	GO: GO:0005198; F:structural molecule activity; IEA. DR	InterPro: IPR000777; GP120. DR	InterPro: IPR007110; Ig-like. DR	Pfam: PF00516; GP120; 1. DR	AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
KW	AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.							
FT	NON TER 1							
FT	NON TER 92							
SQ	SEQUENCE 92 AA; 10482 MW; 3395DB8810C080F1 CRC64;							
	Query Match 96.5%; Score 194; DB 2; Length 92;							
	Best Local Similarity 94.4%; Pred. No. 9e-20; Indels 0; Gaps 0;							
	Matches 34; Conservative 2; Mismatches 0; Indels 0; Gaps 0;							
QY	1 CSRPNNTKRSIPMGPGRAFYYTGGIIGDIRQAHCN 36 : : : : :							
Db	27 CTRPNNTRKSIIPIGPGRAFYYTGGIIGDIRQAHCN 62 : : : : :							
RESULT 9								
Q79067	PRELIMINARY; PRT; 92 AA.							
ID	Q79067 PRELIMINARY; PRT; 92 AA.							
AC	Q79067; (TrEMBLrel. 01, Created)							
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)							
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)							
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)							
DE	Envelope protein (Fragment).							
GN	Name=env;							
OS	Human immunodeficiency virus 1.							
OC	Viruses; Retrovirdae; Retroviridae; Lentivirus.							
OX	NCBI_TaxID=11676;							
RN	[1]							
RP	SEQUENCE FROM N.A.							
RX	MEDLINE=94022315; PubMed=8415653;							
RA	Kuiken C.L., Zwart G., Baan E., Coutinho R.A., Hoek van den J.A.R., Ra Goudamit J.;							

RT "Increasing antigenic and genetic diversity of the HIV-1 V3 domain in
 RL the course of the AIDS epidemic";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:9065-9065(1993).

RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96203966; PubMed=8627267;
 RA Kuiken C.L., Cornelissen M.T.E., Zorgrader F., Hartman S.,
 RA Gibbs A.J., Goudsmit J.;
 RT "Consistent risk group-associated differences in human
 RT immunodeficiency virus type 1 vpr, vpu and V3 sequences despite
 RT independent evolution";
 RL J. Gen. Virol. 77:783-792(1996).

RN [3]
 RP SEQUENCE FROM N.A.
 RA Kuiken L.;
 RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.
 RL EMBL; Z29323; CAA82520.1; --
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000777; GP120.
 DR InterPro; IPR007110; Ig-like.
 DR Pfam; PF00516; GP120; 1.
 DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.

KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
 FT NON_TER 1 92
 FT NON_TER 92 92
 SQ SEQUENCE 92 AA; 10443 MW; DE27D77BDB2283E5 CRC64;

Query Match 96.5%; Score 194; DB 2; Length 92;
 Best Local Similarity 94.4%; Pred. No. 9e-20;
 Matches 34; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHCN 36
 Db 26 CTRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHCN 61

RESULT 10
 Q9DPV9 ID Q9DPV9 PRELIMINARY; PRT; 112 AA.
 AC Q9DPV9;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Envelope glycoprotein (Fragment).
 DE Name=env;
 GN Human immunodeficiency virus 1.
 OS Viruses; Retroviridae; Retroviridae; Lentivirus.
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang Y.M., Saksena N.K.;
 RT "Molecular evidence for drug-induced compartmentalization of HIV-1
 RT quasiespecies in a patient with periodic changes to the HAART
 RT regimens";
 RL AIDS 0:0-0(2001).
 DR EMBL; AF291142; AAC35348.1; --
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.

KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
 FT NON_TER 1 112
 FT NON_TER 112 112
 SQ SEQUENCE 112 AA; 12603 MW; A2F0458A21E5C652 CRC64;

Query Match 96.5%; Score 194; DB 2; Length 112;
 Best Local Similarity 94.4%; Pred. No. 1.1e-19;
 Matches 34; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHCN 36
 Db 26 CTRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHCN 61

RESULT 10
 Q9DPV9 ID Q9DPV9 PRELIMINARY; PRT; 112 AA.
 AC Q9DPV9;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Envelope glycoprotein (Fragment).
 DE Name=env;
 GN Human immunodeficiency virus 1.
 OS Viruses; Retroviridae; Retroviridae; Lentivirus.
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang Y.M., Saksena N.K.;
 RT "Molecular evidence for drug-induced compartmentalization of HIV-1
 RT quasiespecies in a patient with periodic changes to the HAART
 RT regimens";
 RL AIDS 0:0-0(2001).
 DR EMBL; AF291142; AAC35348.1; --
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.

KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
 FT NON_TER 1 112
 FT NON_TER 112 112
 SQ SEQUENCE 112 AA; 12603 MW; A2F0458A21E5C652 CRC64;

Query Match 96.5%; Score 194; DB 2; Length 112;
 Best Local Similarity 94.4%; Pred. No. 1.1e-19;
 Matches 34; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHCN 36
 Db 28 CTRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHCN 63

RESULT 11
 Q9DPW0 ID Q9DPW0 PRELIMINARY; PRT; 112 AA.
 AC Q9DPW0;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Envelope glycoprotein (Fragment).
 DE Name=env;
 GN Human immunodeficiency virus 1.
 OS Viruses; Retroviridae; Retroviridae; Lentivirus.
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang Y.M., Saksena N.K.;
 RT "Molecular evidence for drug-induced compartmentalization of HIV-1
 RT quasiespecies in a patient with periodic changes to the HAART
 RT regimens";
 RL AIDS 0:0-0(2001).
 DR EMBL; AF291141; AAG35347.1; --
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.

KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
 FT NON_TER 1 112
 FT NON_TER 112 112
 SQ SEQUENCE 112 AA; 12587 MW; 16B8F2D91CBB6AEE CRC64;

Query Match 96.5%; Score 194; DB 2; Length 112;
 Best Local Similarity 94.4%; Pred. No. 1.1e-19;
 Matches 34; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHCN 36
 Db 28 CTRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHCN 63

RESULT 12
 Q9DPW1 ID Q9DPW1 PRELIMINARY; PRT; 112 AA.
 AC Q9DPW1;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Envelope glycoprotein (Fragment).
 DE Name=env;
 GN Human immunodeficiency virus 1.
 OS Viruses; Retroviridae; Retroviridae; Lentivirus.
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang Y.M., Saksena N.K.;
 RT "Molecular evidence for drug-induced compartmentalization of HIV-1
 RT quasiespecies in a patient with periodic changes to the HAART
 RT regimens";
 RL AIDS 0:0-0(2001).
 DR EMBL; AF291140; AAG35346.1; --
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.

KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
 FT NON_TER 1 112
 FT NON_TER 112 112
 SQ SEQUENCE 112 AA; 12587 MW; 16B8F2D91CBB6AEE CRC64;

Query Match 96.5%; Score 194; DB 2; Length 112;
 Best Local Similarity 94.4%; Pred. No. 1.1e-19;
 Matches 34; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHCN 36
 Db 28 CTRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHCN 63

RESULT 12
 Q9DPW1 ID Q9DPW1 PRELIMINARY; PRT; 112 AA.
 AC Q9DPW1;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Envelope glycoprotein (Fragment).
 DE Name=env;
 GN Human immunodeficiency virus 1.
 OS Viruses; Retroviridae; Retroviridae; Lentivirus.
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang Y.M., Saksena N.K.;
 RT "Molecular evidence for drug-induced compartmentalization of HIV-1
 RT quasiespecies in a patient with periodic changes to the HAART
 RT regimens";
 RL AIDS 0:0-0(2001).
 DR EMBL; AF291140; AAG35346.1; --
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.

KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
 FT NON_TER 1 112
 FT NON_TER 112 112
 SQ SEQUENCE 112 AA; 12587 MW; 16B8F2D91CBB6AEE CRC64;

Query Match 96.5%; Score 194; DB 2; Length 112;
 Best Local Similarity 94.4%; Pred. No. 1.1e-19;
 Matches 34; Conservative 2; Mismatches 0; Indels 0; Gaps 0;


```
FT NON_TER 112 112
SQ SEQUENCE 112 AA; 12585 MW; A2E8F2D911D6C6E2 CRC64;

Query Match
Best Local Similarity 96.5%; Score 194; DB 2; Length 112;
Matches 34; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHN 36
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 28 CTRPNNTKRSIPGPGRAFYTTGQIIGDIRQAHN 63

RESULT 13
Q9DPW2 PRELIMINARY; PRT; 112 AA.
AC Q9DPW2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN Name-env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RA Wang Y.M., Sakseena N.K.;
RT "Molecular evidence for drug-induced compartmentalization of HIV-1
   quasi-species in a patient with periodic changes to the HAART
   regimens.";
RL AIDS 0:0-(2001).
DR EMBL; AF291139; AAC35345.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 112 112
SQ SEQUENCE 112 AA; 12568 MW; 16F63DE5F1A3FB67 CRC64;

Query Match
Best Local Similarity 96.5%; Score 194; DB 2; Length 112;
Matches 34; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHN 36
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 28 CTRPNNTKRSIPGPGRAFYTTGQIIGDIRQAHN 63

RESULT 14
Q9DPW3 PRELIMINARY; PRT; 112 AA.
AC Q9DPW3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN Name-env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RA Wang Y.M., Sakseena N.K.;
RT "Molecular evidence for drug-induced compartmentalization of HIV-1
   quasi-species in a patient with periodic changes to the HAART
   regimens.";
RL AIDS 0:0-(2001).
DR EMBL; AF291138; AAC35344.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
```

```
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 112 112
SQ SEQUENCE 112 AA; 12586 MW; A2E8F2D91CBB7034 CRC64;

Query Match
Best Local Similarity 96.5%; Score 194; DB 2; Length 112;
Matches 34; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHN 36
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 28 CTRPNNTKRSIPGPGRAFYTTGQIIGDIRQAHN 63

RESULT 15
Q71020 PRELIMINARY; PRT; 113 AA.
AC Q71020;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein, C2-V3 region (Fragment).
GN Name-env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RA MEDLINE; 95194719; PubMed-7888212;
RX Kalish M.L., Luo C.C., Weniger B.G., Limpakarnjanarat K., Young N.,
   Ou C.Y., Schochetman G.;
RT "Early HIV type 1 strains in Thailand were not responsible for the
   current epidemic.";
RL AIDS Res. Hum. Retroviruses 10:1573-1575(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Yealey A.R.;
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U15579; AAC34014.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 113 113
SQ SEQUENCE 113 AA; 12642 MW; B1DBE4BC87669819 CRC64;

Query Match
Best Local Similarity 96.5%; Score 194; DB 2; Length 113;
Matches 34; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHN 36
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 19 CTRPNNTKRSIPGPGRAFYTTGQIIGDIRQAHN 54

Search completed: April 11, 2005, 11:14:04
Job time : 12.5261 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2005, 10:37:31 ; Search time 283.737 Seconds
(without alignments)
1180.438 Million cell updates/sec

Title: US-09-762-261C-1
Perfect score: 4517
Sequence: 1 MRVKGIRNQHGWGCTML.....GRALLHIPTRIOGLERALL 866

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4596	99.5	866	3 AAY79020	Aay79020 HIV-1 env
2	3971.5	86.0	857	2 AAR14904	Aar14904 HIV-1 (MN-
3	3971.5	86.0	857	2 AAW11580	Aaw11580 Human Imm
4	3971.5	86.0	857	2 AAW88112	Aaw88112 Predicted
5	3970	86.0	856	2 AAW11579	Aaw11579 Human Imm
6	3970	86.0	856	2 AAW88111	Aaw88111 Predicted
7	3941.5	85.4	857	8 ADP04371	Adp04371 Envelope
8	3935	85.2	856	2 AAW58805	Aaw58805 HIV-1 iso
9	3933	85.2	856	2 AAR79173	Aar79173 HIV virus
10	3920.5	84.9	883	4 AAB82761	Aab82761 Ancestral
11	3917	84.8	856	2 AAR14903	Aar14903 HIV-1 (MN)
12	3913	84.8	854	2 AAW43068	Aaw43068 HIV-1 gpl
13	3896	84.4	856	6 ABR55495	Abr55495 Amino aci
14	3859.5	83.6	863	8 ADP04369	Adp04369 Envelope
15	3845.5	83.3	847	8 ADK14406	Adk14406 HIV wild-
16	3845.5	83.3	847	8 ADP20081	Adp20081 Human imm
17	3841.5	83.2	847	8 ADK14404	Adk14404 HIV mutan
18	3841	83.2	850	2 AAR67724	Aar67724 gp120 fro
19	3839.5	83.2	847	8 ADK14405	Adk14405 HIV mutan
20	3814	82.6	848	8 ADP20072	Adp20072 Human imm
21	3812.5	82.6	855	2 AAW11581	Aaw11581 Human imm
22	3812.5	82.6	855	2 AAW88113	Aaw88113 Env prote
23	3806.5	82.4	857	2 AAR67725	Aar67725 gp120 fro
24	3795.5	82.2	847	3 AAY97073	Aay97073 Variant H
25	3795.5	82.2	847	8 ADO05103	Ado05103 Human imm

26	3787.5	82.0	851	1 AAP80967	Aap80967 HIV prote
27	3768	81.6	856	2 AAR67726	Aar67726 gp120 fro
28	3758.5	81.4	855	2 AAW43069	Aaw43069 HIV-1 gp1
29	3753.5	81.3	868	5 AAO19389	Aao19389 Lymphaden
30	3744	81.1	856	2 AAR25940	Aar25940 Modified
31	3744	81.1	856	4 AAB85999	Aab85999 Amino aci
32	3744	81.1	863	2 AAR43869	Aar43869 HTLV-III
33	3738.5	81.0	855	8 ADO52541	Ado52541 HIV-1 rec
34	3735.5	80.9	868	1 AAP60422	Aap60422 Sequence
35	3734.5	80.9	868	1 AAP60063	Aap60063 HIV virus
36	3734	80.9	863	1 AAP60349	Aap60349 HTLV-III
37	3733.5	80.9	901	1 AAP70665	Aap70665 Sequence
38	3732.5	80.8	849	8 ADO52531	Ado52531 HIV-1 rec
39	3731	80.8	856	2 AAW89325	Aaw89325 HIV-1 env
40	3731	80.8	856	6 ABU63322	Abu63322 Human lym
41	3731	80.8	856	8 ADO26434	Ado26434 HTLV-III
42	3726.5	80.7	848	5 AAU11872	Aau11872 HIV env p
43	3726	80.7	856	3 AAY97072	Aay97072 Wild type
44	3726	80.7	856	8 ADE84721	Ade84721 Human imm
45	3726	80.7	856	8 ADN36425	Adn36425 HIV prote

ALIGNMENTS

RESULT 1
AAY79020
ID AAY79020 standard; protein; 866 AA.
XX
AC AAY79020;
XX
DT 12-SEP-2003 (revised)
DT 05-JUN-2000 (first entry)
XX
DE HIV-1 envelope protein amino acid sequence.
XX
KW HIV-1; envelope protein; env; cross reactive anti-serum; antibody;
KW vaccine; antibody; detect.
XX
OS Human immunodeficiency virus 1.
XX
PN WO200007631-A1.
XX
PD 17-FEB-2000.
XX
PF 04-AUG-1999; 99WO-US017596.
XX
PR 04-AUG-1998; 98US-0095267P.
XX
PA (JACK-) JACKSON FOUND HENRY M.
XX
PI Quinnan GV, Zhang PF;
XX
DR WPI; 2000-205578/18.
XX
PT Human immunodeficiency virus-1 envelope protein, useful as a vaccine or
PT immunogenic preparation against HIV-1 infection induces the production of
PT broadly cross-reactive neutralizing anti-serum.
XX
PS Claim 2; Page 33; 54pp; English.
XX
CC This sequence represents a human immunodeficiency virus type 1 (HIV-1)
CC envelope (env) protein amino acid sequence. The invention relates to the
CC HIV-1 env protein or its fragments, which when administered to a mammal,
CC induces the production of broadly cross-reactive neutralising anti-serum
CC against multiple strains of HIV-1. The HIV-1 env protein and its
CC fragments are useful as vaccines, immunogenic compositions or diagnostic
CC reagents. They can also be used to detect anti-HIV-1/HIV-1 neutralizing
CC antibodies in a sample. The amino acid sequence and its fragments or mRNA
CC sequence are useful for generating antibodies in a mammal. In addition, a
CC recombinant delivery vector containing the env amino acid sequence may
CC also be used as a vaccine. (Updated on 12-SEP-2003 to standardise OS
CC field)

```
XX SQ Sequence 866 AA;
Query Match 99.5%; Score 4596; DB 3; Length 866;
Best Local Similarity 99.5%; Pred. No. 2.6e-243;
Matches 862; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MRVKGIRRYQHWWGCTMLGLLMICSAATEKLWVTYYGVVPVWKEATTTLCASDAKAY 60
DB 1 MRVKGIRRYQHWWGCTMLGLLMICSAATEKLWVTYYGVVPVWKEATTTLCASDAKAY 60

QY 61 DTEAHNVWATHACVPTDPNPQVELVNVTENFNMKNMVEQMHEDIISLWQSLKPCVK 120
DB 61 DTEAHNVWATHACVPTDPNPQVELVNVTENFNMKNMVEQMHEDIISLWQSLKPCVK 120

QY 121 LTPCLVTLNCTDLRNTNTNNNSSEGTIKGEMKNCNFNATSIIGDKQKEYAL 180
DB 121 LTPCLVTLNCTDLRNTNTNNNSSEGTIKGEMKNCNFNATSIIGDKQKEYAL 180

QY 181 LYKLDIEPIDNDNTSVRLISCNSTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG 240
DB 181 LYKLDIEPIDNDNTSVRLISCNSTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG 240

QY 241 KGSKKNVSTVQCTHGIRPVVSTQLLNGSLAEEVWIRSENFNTNNAKTIIVQLREPVKIN 300
DB 241 KGSKKNVSTVQCTHGIRPVVSTQLLNGSLAEEVWIRSENFNTNNAKTIIVQLREPVKIN 300

QY 301 CSRPNNNTRKSIIPMGPGRAFYTTGIIIGDIRQAHCHNISKNTWTNALKQVVEKLGEOFNKT 360
DB 301 CSRPNNNTRKSIIPMGPGRAFYTTGIIIGDIRQAHCHNISKNTWTNALKQVVEKLGEOFNKT 360

QY 361 KIVFTSSGGDPEIVTHSFNCAGEFFYCNLTQLFDSIWNSENGTWNITRGLNNTGRNDTI 420
DB 361 KIVFTSSGGDPEIVTHSFNCAGEFFYCNLTQLFDSIWNSENGTWNITRGLNNTGRNDTI 420

QY 421 TLPCKRIQIINRQEVGKAMYPPIKGNISCSNITGLLLTRDGGKDDNSRDGNETFRPG 480
DB 421 TLPCKRIQIINRQEVGKAMYPPIKGNISCSNITGLLLTRDGGKDDNSRDGNETFRPG 480

QY 481 GGDMDNRSELKYKVVKIEPLGVAPTAKRVVQREERAVGIGAMFGLGAAGSTMG 540
DB 481 GGDMDNRSELKYKVVKIEPLGVAPTAKRVVQREERAVGIGAMFGLGAAGSTMG 540

QY 541 AASVTLTVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVWGIKQLOARILAVERYLKDQ 600
DB 541 AASVTLTVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVWGIKQLOARILAVERYLKDQ 600

QY 601 LIGTWCSGKLICTTTPVWNASWSKNTLEAIWNNTWQMDKEIDNYTSLIYSLIEESP 660
DB 601 LIGTWCSGKLICTTTPVWNASWSKNTLEAIWNNTWQMDKEIDNYTSLIYSLIEESP 660

QY 661 IQEKNQOELELDKWANLWDFISNWLWYIKIFIMVGLVGLRIVFVVLISVNRVQ 720
DB 661 IQEKNQOELELDKWANLWDFISNWLWYIKIFIMVGLVGLRIVFVVLISVNRVQ 720

QY 721 GYSPLSFQTRLPAPRGDPREEIEEGDRDRSGLLVDGFTLIWDLRSCLFSYHR 780
DB 721 GYSPLSFQTRLPAPRGDPREEIEEGDRDRSGLLVDGFTLIWDLRSCLFSYHR 780

QY 781 LRDLILLIVTRIVELLGRGWEILLKYWNLLQYWSQELKNSAVSIFNATAVAEGTDRVI 840
DB 781 LRDLILLIVTRIVELLGRGWEILLKYWNLLQYWSQELKNSAVSIFNATAVAEGTDRVI 840

QY 841 QVLQVGRALLHIPTRIQGLERALL 866
DB 841 EVLQVGRALLHIPTRIQGLERALL 866

RESULT 2
AAR14904
ID AAR14904 standard; protein; 857 AA.
XX
AC AAR14904;
```

```
XX 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 17-DEC-2001 (revised)
DT 05-FEB-1992 (first entry)
XX HIV-1 (MN-ST1) env protein.
XX human immunodeficiency virus; United States; MN isolate; AIDS;
KW envelope protein.
XX Human immunodeficiency virus 1.
XX USN7599491-N.
XX 15-OCT-1991.
XX 17-OCT-1990; 90US-00183830.
XX 17-OCT-1990; 90US-00599491.
XX (USSH ) NAT INST OF HEALTH.
XX Reitz M;
XX WPI; 1991-346752/47.
XX P-PSDB; AAR14904.
XX US HIV-1 isolates MN-ST1 and BA-L, ENV protein and DNA - are useful in
PT therapeutics, vaccines and diagnostic tests.
XX Example 2; Fig 6; 61pp; English.
XX The infectious molecular clone, lambda MN-ST1, was obtained by cloning
CC integrated provirus from DNA purified from peripheral blood lymphocytes
CC infected with HIV-1 (MN) and maintained in culture for one month. The
CC integrated proviral DNA was partially digested with Sau3A to give
CC fragments of 15-20 kb. The fragments were cloned in EMBL3 and the entire
CC sequence of the clone was determined and the amino acid sequence of the
CC env protein was deduced from it. (Note: Revised entry submitted to
CC correct the patent number format of US Government-owned NTIS applications
CC to prevent clashes with ongoing US granted patent numbers. For further
CC information please visit the Derwent web site at
CC www.derwent.com/dwpi/updates/ntis.us.html.) (Updated on 25-MAR-2003 to
CC correct PR field.) (Updated on 27-AUG-2003 to correct OS field.)
XX Sequence 857 AA;
SQ
Query Match 86.0%; Score 3971.5; DB 2; Length 857;
Best Local Similarity 86.6%; Pred. No. 3.7e-209;
Matches 750; Conservative 37; Mismatches 70; Indels 9; Gaps 5;

QY 1 MRVKGIRRYQHWWGCTMLGLLMICSAATEKLWVTYYGVVPVWKEATTTLCASDAKAY 60
DB 1 MRVKGIRRYQHWWGCTMLGLLMICSAATEKLWVTYYGVVPVWKEATTTLCASDAKAY 60

QY 61 DTEAHNVWATHACVPTDPNPQVELVNVTENFNMKNMVEQMHEDIISLWQSLKPCVK 120
DB 61 DTEAHNVWATHACVPTDPNPQVELVNVTENFNMKNMVEQMHEDIISLWQSLKPCVK 120

QY 121 LTPCLVTLNCTDLRNTNTNNNSSEGTIKGEMKNCNFNATSIIGDKQKEYAL 180
DB 121 LTPCLVTLNCTDLRNTNTNNNSSEGTIKGEMKNCNFNATSIIGDKQKEYAL 180

QY 181 LYKLDIEPIDNDNTSVRLISCNSTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG 240
DB 181 LYKLDIEPIDNDNTSVRLISCNSTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG 240

QY 241 KGSKKNVSTVQCTHGIRPVVSTQLLNGSLAEEVWIRSENFNTNNAKTIIVQLREPVKIN 300
DB 241 KGSKKNVSTVQCTHGIRPVVSTQLLNGSLAEEVWIRSENFNTNNAKTIIVQLREPVKIN 300

QY 301 CSRPNNNTRKSIIPMGPGRAFYTTGIIIGDIRQAHCHNISKNTWTNALKQVVEKLGEOFNKT 360
```

```

Db      301 CTRPNYKRIHIGPGRAFYYTNNIIIGTIRQAHCHNISRAKWNLTLEQIVSKLKEQFKNK 360
Qy      361 KIVFTNSSGGDPEIVTHSFNCAGFEFFYCNTTQOLFDSIWNSENGTWNITRGLNNTGRNDTI 420
Db      361 TIVFNQSSGGDPEIVMHSFNCAGFEFFYCNTSPLFNSTWNG-NNTWN-----NTTGSNNI 414
Qy      421 TLPCKRIKOIINRWQEVCKAMYPAPIKNISCSSNITGLLITRDGKDDNSRDGNETFRPG 480
Db      415 TLQCKIKOIINRWQEVCKAMYPAPIEGQIRCCSSNITGLLITRDGKDDTND-TEIFRPG 473
Qy      481 GGDNRDNRSELYKYKVKIEPLGVAPTAKRRVVQREERAVGLGAMFFGFLGAAGSTM 540
Db      474 GGDNRDNRSELYKYKVKITIEPLGVAPTAKRRVVQREKRA-AIGALFLGLGAGSTM 532
Qy      541 AASVTLTVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVWGIKOLQARILAVERYLKDQ 600
Db      533 AASVTLTVQARQLLSGIVQOQSNLLRAIEAQHMLQLTVWGIKOLQARILAVERYLKDQ 592
Qy      601 LLGIWCSGKLICTTTVPWNASWSKNTLEAIWNNTWQWDKEIDNYTSLIYSLIESP 660
Db      593 LLGIWCSGKLICTTTVPWNASWS-NKSLDDIWNNTWQWDKEIDNYTSLIYSLIESQ 651
Qy      661 IQEKNQELLEDKQWNLWNPDISNWLWYIKIFIMIVGLVGLRIVFVVLVSVNRVQ 720
Db      652 TQEMNEQELLELDKQWNLWNPDISNWLWYIKIFIMIVGLVGLRIVFVVLVSVNRVQ 711
Qy      721 GYSPLSQTRLPAPRGDRPEIEEGGDRDRDRSGLLVDGFLTLIIVWDLRSLCLFSYHR 780
Db      712 GYSPLSLQTRPPVPRGDRPEIEEGGDRDRDRSGRLVHGLAIIVWDLRSLFLSYHR 771
Qy      781 LRDLIIIVTRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATATAVAGETDRI 840
Db      772 LRDLIIIAARIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATATAVAGETDRI 831
Qy      841 QVLQVRGALLHIPTRIQGLERALL 866
Db      832 EVLQVRGALLHIPTRIQGLERALL 857

RESULT 3
AAW11580
ID AAW11580 standard; protein; 857 AA.
XX
AC AAW11580;
XX
XX 17-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 25-MAR-1997 (first entry)
XX
XX Human Immunodeficiency Virus-1 strain MN-ST1 envelope protein.
XX
XX Acquired immune deficiency syndrome; AIDS; envelope protein; env gene;
XX vaccine.
XX
XX Human immunodeficiency virus 1; (strain MN-ST1).
XX
XX US5576000-A.
XX
XX 19-NOV-1996.
XX
XX 15-FEB-1995; 95US-00388809.
XX
XX 17-OCT-1990; 90US-00599491.
XX
XX 25-FEB-1993; 93US-00022835.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Gallo RC, Franchini G, Lori FC, Popovic M, Reitz MS, Gartner S;
XX Markham PD;
XX
XX WPI; 1997-011206/01.
XX
XX N-PSDB; AAT58550.

```

```

XX
PT New isolated envelope protein of HIV-1 strain BA-L and recombinant
PT equivalents - useful as immunogens for vaccines and antibody prodn.,
XX typical of us clinical isolates.
PS Example; Fig 6; 86pp; English.
XX
CC The infectious molecular clone, lambda MN-ST1, was obtained by cloning
CC integrated provirus from DNA purified from peripheral blood lymphocytes
CC infected with HIV-1 (MN) and maintained in culture for one month. The
CC integrated proviral DNA was partially digested with Sau3A to produce
CC fragments of 15-20 kb. The fragments were cloned into the BamHI site of
CC lambda EMBL3. The DNA sequence of the entire clone was determined and the
CC env gene was excised and cloned into an expression plasmid for
CC recombinant production of the envelope protein. The env protein has the
CC present sequence and can be used as an immunogen for raising antibodies
CC against HIV. (Updated on 25-MAR-2003 to correct PF field.) (Updated on 17
CC -OCT-2003 to standardise OS field)
XX
SQ Sequence 857 AA;
Query Match 86.0%; Score 3971.5; DB 2; Length 857;
Best Local Similarity 86.6%; Pred. No. 3.7e-209;
Matches 750; Conservative 37; Mismatches 70; Indels 9; Gaps 5;
Qy 1 MRVKGIRRNQHWGWTMLLGLLMTCSAATEKLWVTYVYGVVPWKEATTLFCASDAKAY 60
Db 1 MRVKGIRRNQHWGWTMLLGLLMTCSAATEKLWVTYVYGVVPWKEATTLFCASDAKAY 60
Qy 61 DTEAHNVWATHACVPTDPNPQEVLEVNVNTEFNKNMKNMVEQMHEDIISLDQSLKPCVK 120
Db 61 DTEAHNVWATHACVPTDPNPQEVLEVNVNTEFNKNMKNMVEQMHEDIISLDQSLKPCVK 120
Qy 121 LTPCLVTLNCTDLRNTNTNTNNSDNNNSSEGIKGGEMKNCSEFNATISGDKMKQKEYAL 180
Db 121 LTPCLVTLNCTDLRNTNTNTNNSDNNNSSEGIKGGEMKNCSEFNATISGDKMKQKEYAL 180
Qy 181 LYKLDIEPDNDNTSYRLISCNSTSVITQACPKISFBPIPIHYCAPAGFAILKCNCKPFG 240
Db 181 LYKLDIVSINNDSTSYRLISCNSTSVITQACPKISFBPIPIHYCAPAGFAILKCNCKPFG 240
Qy 241 KGSCKNVSTVQCTHGIRPVVSTOLLNGSLAEVEVIRSENFNTNNAKTIIVQLREPKIN 300
Db 241 KGSCKNVSTVQCTHGIRPVVSTOLLNGSLAEVEVIRSENFNTNNAKTIIVHLNESVQIN 300
Qy 301 CSRPNNTKRSIPMGPCRAFYTTGQIIGIRQAHCHNISKTNTNNAKQVVEKGEQFNKT 360
Db 301 CTRPNYKRIHIGPGRAFYYTNNIIIGTIRQAHCHNISRAKWNLTLEQIVSKLKEQFKNK 360
Qy 361 KIVFTNSSGGDPEIVTHSFNCAGFEFFYCNTTQOLFDSIWNSENGTWNITRGLNNTGRNDTI 420
Db 361 TIVFNQSSGGDPEIVMHSFNCAGFEFFYCNTSPLFNSTWNG-NNTWN-----NTTGSNNI 414
Qy 421 TLPCKRIKOIINRWQEVCKAMYPAPIKNISCSSNITGLLITRDGKDDNSRDGNETFRPG 480
Db 415 TLQCKIKOIINRWQEVCKAMYPAPIEGQIRCCSSNITGLLITRDGKDDTND-TEIFRPG 473
Qy 481 GGDNRDNRSELYKYKVKIEPLGVAPTAKRRVVQREERAVGLGAMFFGFLGAAGSTM 540
Db 474 GGDNRDNRSELYKYKVKITIEPLGVAPTAKRRVVQREKRA-AIGALFLGLGAGSTM 532
Qy 541 AASVTLTVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVWGIKOLQARILAVERYLKDQ 600
Db 533 AASVTLTVQARQLLSGIVQOQSNLLRAIEAQHMLQLTVWGIKOLQARILAVERYLKDQ 592
Qy 601 LLGIWCSGKLICTTTVPWNASWSKNTLEAIWNNTWQWDKEIDNYTSLIYSLIESP 660
Db 593 LLGIWCSGKLICTTTVPWNASWS-NKSLDDIWNNTWQWDKEIDNYTSLIYSLIESQ 651
Qy 661 IQEKNQELLEDKQWNLWNPDISNWLWYIKIFIMIVGLVGLRIVFVVLVSVNRVQ 720
Db 652 TQEMNEQELLELDKQWNLWNPDISNWLWYIKIFIMIVGLVGLRIVFVVLVSVNRVQ 711

```


XX Acquired immune deficiency syndrome; AIDS; envelope protein; env gene;
 KW vaccine.
 XX Human immunodeficiency virus 1; (strain MN-PH1).
 OS
 XX US5576000-A.
 PN
 XX 19-NOV-1996.
 PD
 XX
 XX 15-FEB-1995; 95US-00388809.
 XX
 XX 17-OCT-1990; 90US-00599491.
 PR
 XX 25-FEB-1993; 93US-00022835.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA
 XX Gallo RC, Franchini G, Lori FC, Popovic M, Reitz MS, Gartner S;
 PI Markham PD;
 PI
 XX WPI; 1997-011206/01.
 DR N-PSDB; AAT58549.
 XX
 XX New isolated envelope protein of HIV-1 strain BA-L and recombinant
 PT equivalents - useful as immunogens for vaccines and antibody prodn.,
 PT typical of US clinical isolates.
 PT
 XX Example; Fig 3; 86pp; English.
 XX
 XX The permuted circular unintegrated viral DNA representing the complete
 CC HIV-1(MN) genome was cloned into the EcoRI site of lambda gtWES.lambda B
 CC DNA from total DNA of H9 cells producing HIV-1(MN) isolate. The clone was
 CC designated lambda MN-PH1. The present sequence was deduced from the env
 CC gene region of the MN-PH1 genome. (Updated on 25-MAR-2003 to correct PF
 CC field.) (Updated on 17-OCT-2003 to standardise OS field)
 CC
 XX Sequence 856 AA;
 SQ

Query Match 86.0%; Score 3970; DB 2; Length 856;
 Best Local Similarity 86.7%; Pred. No. 4.5e-209;
 Matches 751; Conservative 36; Mismatches 69; Indels 10; Gaps 6;
 1 MRVKGIRRYQHWWGWTMLGLLMTCATEKLWVTYGVVPVWKEATTLFCASDAKAY 60
 1 MRVKGIRRYQHWWGWTMLGLLMTCATEKLWVTYGVVPVWKEATTLFCASDAKAY 60
 61 DTEAHNVWATHACVPTDPNPQVELVNVTENFNWKNMVEQMHEDIISLWDQSLKPCVK 120
 61 DTEAHNVWATHACVPTDPNPQVELVNVTENFNWKNMVEQMHEDIISLWDQSLKPCVK 120
 121 LTPCLVTNLCTDLRNTTNTNNSNSSEGTIKGEMKNCSEFNATSIIGDRMQKEYAL 180
 121 LTPCLVTNLCTDLRNTTNTNNSNSSEGTIKGEMKNCSEFNATSIIGDRMQKEYAL 180
 181 LYKLDIEPDNDNTSLRISNTSVITQACPKISFEPPIHYCAPAGFALLKNDKFFSG 240
 181 LYKLDIVSIDNSTYRLISNTSVITQACPKISFEPPIHYCAPAGFALLKNDKFFSG 240
 241 KGSCKNVSTVQCHGIRPVVSTOLLNGSLAEVWIRSENFTNNAKTIIVQLREPVKN 300
 241 KGSCKNVSTVQCHGIRPVVSTOLLNGSLAEVWIRSENFTNNAKTIIVHLNVSQIN 300
 301 CSRPNNNTKSIIPWGPGRAPYTTGQIIGDIRQAHNCISKTNNTNALKQVVEKLGEOFNKT 360
 301 CTFPNYKRRKIHIGPGRAPYTTNIGTIRQAHNCISRAKNDTLRQIVSKLKEQFNK 360
 361 KIVFTNSSGGDPRIVTHSFNCAGEFFVNTQTFDSIWNSENCTWMTIRGLNNTGRNDTI 420
 361 TIVFNQSSGGDPRIVHVSFNCGGEFFVNTSPLFNSTWNG-NNTWN-----NTTGSNNNI 414
 421 TLPCKRIKQIINRQVECKKAMYPPIKGNISCSNITGLLITRGGKVDNSRDGNFTFRG 480
 415 TLOCKIKQIINRQVECKKAMYPPIEQIRCSNITGLLITRGGKVDNTD-TEIFRPG 473

QY 481 GGDMDRNWSELYKYYKVVIEPLGVAPTAKARRVQREERAVGLGAMFFGFLGAAGSTMG 540
 DB 474 GGDMDRNWSELYKYYKVVIEPLGVAPTAKARRVQREKRA-AIGALFLGFLGAAGSTMG 532
 QY 541 AASVTLTVQARQLLSGIVOOOSNLLRAIEAQOHLLOLTVMGIKOLOARILAVERYLKDQO 600
 DB 533 AASVTLTVQARQLLSGIVOOONLLRAIEAQOHLLOLTVMGIKOLOARVLAVERYLKDQO 592
 QY 601 LLGIMGCSGKLICTTTVPWNASMSKNTLEAIWNNMTWQMDKEIDNYTSLIYSLIBESP 660
 DB 593 LLGFWGCSGKLICTTTVPWNASMS-NKSLDDIWNNTWQWEREIDNYTSLIYSLLEKSQ 651
 QY 661 IQOEKNEQELLELDKWANLWNNFIDINWLMYIKIFIMIVGGLVGLRIVFVVLIVNVRVQ 720
 DB 652 TQOEKNEQELLELDKWASLWNNFIDITNWLWYIKIFIMIVGGLVGLRIVFAVLIVNVRVQ 711
 QY 721 GYSPLSFQTRLPAPRGPDRPEETEEEGGDRDRSGLLVDGPTLTLLIWDLRSICLFSYHR 780
 DB 712 GYSPLSLQTRPPVPRGPDRPEETEEEGGDRDRSGRLVHGFLAIIWVDLRSILFSYHH 771
 QY 781 LRDLLIVTRIVELLGRRGWEILKYWNLLQYWSQELKNSAVSLFNATAIYVAEGTDRVI 840
 DB 772 -RDLIIAARIVELLGRRGWEVLKYWNLLQYWSQELKNSAVSLNATAIYVAEGTDRVI 830
 QY 841 QVLQVRGALLHIPTRIQGLERALL 866
 DB 831 EVLQVRGAILHIPTRIQGLERALL 856

RESULT 6
 AAW88111
 ID AAW88111 standard; protein; 856 AA.
 XX
 AC AAW88111;
 XX
 DT 17-OCT-2003 (revised)
 DT 09-APR-1999 (first entry)
 XX
 DE Predicted amino acid sequence of the MN-PH1 envelope (env) protein.
 XX
 XX MN-PH1 clone; HIV-1; HIV-1 strain BA-L; envelope protein; vaccine;
 KW immunotherapy; HIV infection; immunogen; HIV-1 diagnosis; env; ds.
 XX
 OS Human immunodeficiency virus 1.
 XX
 PN US5869313-A.
 XX
 PD 09-FEB-1999.
 XX
 XX 14-MAY-1996; 96US-00647714.
 XX
 XX 17-OCT-1990; 90US-00599491.
 PR 25-FEB-1993; 93US-00022835.
 PR 15-FEB-1995; 95US-00388809.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA
 XX Gallo RC, Franchini G, Reitz MS, Lori FC, Popovic M, Gartner S;
 PI Markham PD;
 PI
 XX WPI; 1999-152779/13.
 DR
 XX DNA encoding env protein of the human immune deficiency virus isolate BA-
 PT L - useful for producing protein for use in vaccines, as assay reagent
 PT and to generate antibodies.
 XX
 PS Example 1; Fig 3A-C; 87pp; English.
 XX
 CC The present sequence represents the predicted amino acid sequence of the
 CC MN-PH1 clone envelope (env) protein. MN-PH1 is a Human immunodeficiency
 CC virus type 1 (HIV-1) clone. The specification also describes the env and
 CC rev coding sequences of the HIV-1 strain BA-L (ATCC 40890). BA-L is more

CC typical of United States isolates of HIV-1 than previously known strains.
CC Recombinant, complete env protein of the BA-L strain is used as a vaccine
CC component and for immunotherapy of existing HIV infections, to detect HIV
CC -specific antibodies, e.g. in donated blood, and as an immunogen to raise
CC specific antibodies, for HIV-1 diagnosis. (Updated on 17-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 856 AA;

Query Match 86.0%; Score 3970; DB 2; Length 856;
Best Local Similarity 86.7%; Pred. No. 4.5e-209;
Matches 751; Conservative 36; Mismatches 69; Indels 10; Gaps 6;
QY 1 MRVKGIRRYQHWWGTMLLGLLMICSAATEKLWVTYYGVVPVWKEATTTLCASDAKAY 60
DB 1 MRVKGIRRYQHWWGTMLLGLLMICSAATEKLWVTYYGVVPVWKEATTTLCASDAKAY 60
QY 61 DTEAHNVWATHACVPTDPNPQVELVNVTFENFMNKNMVEQMHEDIISLWDSLKPCVK 120
DB 61 DTEAHNVWATHACVPTDPNPQVELVNVTFENFMNKNMVEQMHEDIISLWDSLKPCVK 120
QY 121 LTPCLVTLNCTDLRNTNTNNNSSEGTIKGEMKNCNFNATISIGDKWQKEYAL 180
DB 121 LTPCLVTLNCTDLRNTNTNNNSSEGTIKGEMKNCNFNATISIGDKWQKEYAL 180
QY 181 LYKLDIEPDNDNTSYRLISCNSTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG 240
DB 181 LYKLDIVSDNSTSYRLISCNSTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG 240
QY 241 KGSCKNVSTVQCTHGRPVVSTOLLNGSLAEVEVIRSENFNTNNAKTIIVQUREPVKIN 300
DB 241 KGSCKNVSTVQCTHGRPVVSTOLLNGSLAEVEVIRSENFNTNNAKTIIVHLNESVQIN 300
QY 301 CSPNNNTRKSIWPGCRAPYTTGIIIGDTRQACHNISKTNTNALKOVVEKLGEOFNKT 360
DB 301 CTSPNNKRIHIGCRAPYTTGIIIGDTRQACHNISRAKMWDTLQIVSKLKEQFKKK 360
QY 361 KIVFTNSSGGDDPIVTHSFNCAGEFFYCNFTTOLFDSIWNSENGTWNITRGLNNTGENDTI 420
DB 361 TIVFNQSSGDDPIVHVSFNCGEFFYCNFTSPLFNSTWNG-NNTWN----NTTGNNNI 414
QY 421 TLPCTRKOIINRWQEVGKAWAPPIKGNITSCSNITGLLTRGGKDNDRGNETFRPG 480
DB 415 TLQCKIKOIINRWQEVGKAWAPPIEQICRCSNITGLLTRGGKDTDND-TEIFRFG 473
QY 481 GGDMDNRWSELYKVKIPELGVAPTAKRRVQREBRVGLGAMFGLGAAGSTWG 540
DB 474 GGDMDNRWSELYKVKIPELGVAPTAKRRVQREKRA-AIGALFLGFLGAAGSTWG 532
QY 541 AASVTLTVQARQLLSGIVQQSNLLRAIEAQHLLQLTVWGIKQLQARILAVERYLKDQ 600
DB 533 AASVTLTVQARQLLSGIVQQSNLLRAIEAQHMLQLTVWGIKQLQARVLAVERYLKDQ 592
QY 601 LLGIWCSGKLICTTVPWNASHSKNTLEA IWNNTWQWDEIKNYTSLSYLSIEESP 660
DB 593 LLGFWCSGKLICTTVPWNASHS-NKSLDDIWNNTWQWDEIKNYTSLSYLSLEKSO 651
QY 661 IQEKNQEQLLEDKWNALWNPDIENWLYIKIFTMIVGGLVGLRIVFVLSIVNRVQ 720
DB 652 TOEKNEQELLELDKWSLWNPDIENWLYIKIFTMIVGGLVGLRIVFVLSIVNRVQ 711
QY 721 GYSPLSFQTLRPPRPEEIEEGGDRDRSGLLVGFTLLIWDLSRLSCLFSYHR 780
DB 712 GYSPLSLQTRPPRPPRPEEIEEGGDRDRSGLRVHGFLLAIIVDLSRLSFLFSYH 771
QY 781 LRDLILLIVTRIVELLGRGWEILLKYWNLLQYWSQELKNSAVSLFNATAIAVAGTDRVI 840
DB 772 -RDLLIAAIVELLGRGWEVLKYWNLLQYWSQELKSSAVSLNATAIAVAGTDRVI 830
QY 841 QVLQVRGALLHPTIRIOGLERALL 866
DB 831 EVLQVRGAILHPTIRIOGLERALL 856

RESULT 7
ADP04371
ID ADP04371 standard; protein; 857 AA.
XX AC ADP04371;
XX DT 26-AUG-2004 (first entry)
XX DE Envelope protein gp120 from antibody sensitive HIV strain.
XX KW anti-HIV; vaccine; human immunodeficiency virus type-1;
KW envelope glycoprotein; cross-reactive immune response; gp41; CD4-binding;
KW co-receptor-binding; outer domain; gp120.
XX OS Human immunodeficiency virus 1.
XX FN WO2004045495-A2.
XX PD 03-JUN-2004.
XX PF 12-MAY-2003; 2003WO-US014721.
XX PR 10-MAY-2002; 2002US-0379052P.
XX (JACK-) JACKSON FOUND HENRY M.
XX Quinman GV;
XX WPI; 2004-431796/40.
XX N-PSDB; ADP04370.
XX Identifying an HIV-1 envelope protein that produces a cross-reactive
PT immune response in a mammal, comprises introducing amino acid
PT substitutions into gp120, and identifying substitutions that produce a
PT cross-reactive immune response.
XX Disclosure; SEQ ID NO 4; 68pp; English.
XX The invention relates to a method of identifying a human immunodeficiency
CC virus type-1 (HIV-1) envelope protein which produces a cross-reactive
CC immune response following administration in a mammal by: (a) substituting
CC one or more amino acids in or near the gp41, CD4-binding, co-receptor-
CC binding or outer domains of gp120; and (b) identifying one or more amino
CC acid substitutions in the domains that produce a cross-reactive immune
CC response following administration in a mammal. The method is useful for
CC identifying an HIV-1 envelope protein, which produces a cross-reactive
CC immune response following administration in a mammal. Compositions
CC comprising such HIV-1 envelope protein may be used as HIV subunit vaccine
CC or immunogenic composition that can induce antibodies that neutralize HIV
CC strains across different phenotypes and clades, or for the prevention or
CC therapy of HIV-1 infection. This sequence corresponds to the gp120
CC envelope glycoprotein from a neutralisation antibody-sensitive (T-cell
CC line adapted) MN strain of HIV-1 (MN-P).
XX SQ Sequence 857 AA;

Query Match 85.4%; Score 3941.5; DB 8; Length 857;
Best Local Similarity 85.7%; Pred. No. 1.6e-207;
Matches 742; Conservative 46; Mismatches 69; Indels 9; Gaps 5;
QY 1 MRVKGIRRYQHWWGTMLLGLLMICSAATEKLWVTYYGVVPVWKEATTTLCASDAKAY 60
DB 1 MRVKGIRRYQHWWGTMLLGLLMICSAATEKLWVTYYGVVPVWKEATTTLCASDAKAY 60
QY 61 DTEAHNVWATHACVPTDPNPQVELVNVTFENFMNKNMVEQMHEDIISLWDSLKPCVK 120
DB 61 DTEAHNVWATHACVPTDPNPQVELVNVTFENFMNKNMVEQMHEDIISLWDSLKPCVK 120
QY 121 LTPCLVTLNCTDLRNTNTNNNSSEGTIKGEMKNCNFNATISIGDKWQKEYAL 180
DB 121 LTPCLVTLNCTDLRNTNTNNNSSEGTIKGEMKNCNFNATISIGDKWQKEYAL 180

QY 181 LYKLDIEPIDNDNTSYRLISNTSVITQACPKISFEPIPIHYCAPAGFAILKCNCKFSG 240
 Db 181 LYKLDIEPIDNDNTSYRLISNTSVITQACPKISFEPIPIHYCAPAGFAILKCNCKFSG 240
 QY 241 KGSCKNVSTVQCTHGIRPVVSTOLLNGSLAEBEVIRSENFNNAKTIIIVQLREPVKIN 300
 Db 241 KGSCKNVSTVQCTHGIRPVVSTOLLNGSLAEBEVIRSENFNNAKTIIIVHLKESVQIN 300
 QY 301 CSRPNNNRKSIIPMGPGRAFPTTGGIIGDIRQAHNCISKTNNWNAKQVVEKLGEOFNKT 360
 Db 301 CTRPNYNKRRIIHIGPGRAFPTTGNIGTIRQAHCTISRAKWNDDTLRQIVSKLKEQFNK 360
 QY 361 KIYFTNSSGGDPEIVTHSFNCAGEFFVCNTTOLFDSIWNSENGTWNITRGLNNTGRDNTI 420
 Db 361 TIYFNPSGGDPEIVMHSFNCAGEFFVCNTTOLFDSIWNSENGTWNITRGLNNTGRDNTI 420
 QY 421 TLPCKRIQIINRWQEVGKAMYAPPIKGNISCSNITGLLTRDGGKDNDRGNETFRPG 480
 Db 415 TLOCKIKIINRWQEVGKAMYAPPIEQIRCSNITGLLTRDGGEDTDND-TEIFRPG 473
 QY 481 GGDMDNRSELYKYKVVKIEPLGVAPTAKRRVVOREERAVGLGAMFFGFLGAAGSTM 540
 Db 474 GGDMDNRSELYKYKVVVTEPLGVAPTAKRRVVOREKRA-AIGALFLGFLGAAGSTM 532
 QY 541 AASVTLTVQAROLLGSGIVQOQNLLRAIEAQOHLIQTWNGIKOLOARILAVERYLKDOQ 600
 Db 533 AASVTLTVQAROLLGSGIVQOQNLLRAIEAQOQNMLQTVWGIKOLQARVOAVERYLKDOQ 592
 QY 601 LLGWGSGKGLICTTTPVWNASKNKTLEAIWNNMTWMDKDEIDNTSLIYSLIEESP 660
 Db 593 LLGWGSGKGLICTTTPVWNASWS-NKSLDDIWNMTWMDKDEIDNTSLIYSLIEKSO 651
 QY 661 IQQEKNEQELLELDKWNANWFDISNWLWYIKIFIMVGLVGLRLIVFVVLISVNRVQ 720
 Db 652 TQQEKNEQELLELDKWNANWFDITNWLWYIKIFIMVGLVGLRLIVFVVLISVNRVQ 711
 QY 721 GYSPLSQTLAPRGPDRPEEBEEDGDRDRSGLLVGDLPLTLIWDLRSLCLFSYHR 780
 Db 712 GYSPLSQTLAPRGPDRPEEBEEDGDRDRSGLLVHGFLLAIWDLRLSLFLSYHH 771
 QY 781 LRDLIIIVTRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATATAVAGETDRVI 840
 Db 772 LRDLIIIVTRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATATAVAGETDRVI 831
 QY 841 QVLQVGRALLHIPTRIQGLERALL 866
 Db 832 EVLQVGRALLHIPTRIQGLERALL 857
 RESULT 8
 AAW58805
 ID AAW58805 standard; protein; 856 AA.
 XX
 AC AAW58805;
 XX
 DT 17-OCT-2003 (revised)
 DT 23-JUL-1998 (first entry)
 XX
 DE HIV-1 isolate MN envelope protein gp120.
 XX
 KW Envelope protein; gp120; peptomer; peptides oligomer; detection;
 KW CD4-binding region; anti-HIV antibody; therapeutic; infection;
 KW inhibition; T cell; immunogenic agent; antibody; vaccine; epitope.
 XX
 OS Human immunodeficiency virus 1.
 XX
 Key Location/Qualifiers
 FT Region 301..335
 FT /label= V3 loop
 FT Region 401..800
 FT /label= CD4-binding_region
 FT Region 801..855
 FT /note= "region with MHC II homology"

XX US5750332-A.
 PN 12-MAY-1998.
 PD 19-JAN-1995; 95US-00375100.
 PF 19-JAN-1994; 94US-00184330.
 PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Robert-Guroff M, Robey FA, Harris-Kelson TA,
 PI WFI; 1998-296756/26.
 PS HIV gp120 peptide oligomers - useful as diagnostic, immunogenic or
 PT therapeutic agents.
 PT Disclosure; Fig 1; 25pp; English.
 PS This sequence represents the Human Immunodeficiency Virus Type I (HIV-I)
 CC isolate MN envelope protein gp120. This protein is used in a novel method
 CC to produce and analyse peptides oligomers, "peptomers", consisting of
 CC several peptide monomers arranged head to tail, each monomer comprising a
 CC sequence from a CD4-binding region of HIV gp120. Such peptides are useful
 CC as assay reagents for detecting anti-HIV antibodies in biological
 CC samples, e.g. blood samples, or as a therapeutic agent for inhibiting
 CC infection of T cells by HIV and as an immunogenic agent for antibody
 CC production e.g. in vaccines. They are preferably synthetically produced
 CC and thus are consistent in their sequence and structure. Most HIV
 CC vaccines and detection peptides used currently rely on fragments of
 CC inactivated live viruses and thus may not contain the correct epitopes to
 CC elicit an immune response. (Updated on 17-OCT-2003 to standardise OS
 CC field)
 XX Sequence 856 AA;
 SQ
 Query Match 85.2%; Score 3935; DB 2; Length 856;
 Best Local Similarity 86.1%; Pred. No. 3.7e-207;
 Matches 746; Conservative 36; Mismatches 74; Indels 10; Gaps 6;
 QY 1 MRVKGIRRNQHWGWTMLGLLMICSAATEKLWTVVYGVVWKEATTTLFCASDAKAY 60
 Db 1 MRVKGIRRNQHWGWTMLGLLMICSAATEKLWTVVYGVVWKEATTTLFCASDAKAY 60
 QY 61 DTEAHNVWATHACVPTDPNPQVEVELNVNTEFNKNMKNMVEQMHEDIISLDQSLKPCVK 120
 Db 61 DTEAHNVWATHACVPTDPNPQVEVELNVNTEFNKNMKNMVEQMHEDIISLDQSLKPCVK 120
 QY 121 LTPLCVTLNCTDLRNTTNTNNSTNNNSSEGTIKGEMKNCSEFNATISIDKMKQKEYAL 180
 Db 121 LTPLCVTLNCTDLRNTTNTNNSTNNNSSEGTIKGEMKNCSEFNATISIDKMKQKEYAL 180
 QY 181 LYKLDIEPIDNDNTSYRLISNTSVITQACPKISFEPIPIHYCAPAGFAILKCNCKFSG 240
 Db 181 LVKLDIVPIDNDNTSYRLISNTSVITQACPKISFEPIPIHYCAPAGFAILKCNCKFSG 240
 QY 241 KGSCKNVSTVQCTHGIRPVVSTOLLNGSLAEBEVIRSENFNNAKTIIIVQLREPVKIN 300
 Db 241 KGSCKNVSTVQCTHGIRPVVSTOLLNGSLAEBEVIRSENFNNAKTIIIVHLNESVQIN 300
 QY 301 CSRPNNNRKSIIPMGPGRAFPTTGGIIGDIRQAHNCISKTNNWNAKQVVEKLGEOFNKT 360
 Db 301 CTRPNYNKRRIIHIGPGRAFPTTGNIGTIRQAHNCISKTNNWNAKQVVEKLGEOFNKT 360
 QY 361 KIYFTNSSGGDPEIVTHSFNCAGEFFVCNTTOLFDSIWNSENGTWNITRGLNNTGRDNTI 420
 Db 361 TIYFNPSGGDPEIVMHSFNCAGEFFVCNTTOLFDSIWNSENGTWNITRGLNNTGRDNTI 420
 QY 421 TLPCKRIQIINRWQEVGKAMYAPPIKGNISCSNITGLLTRDGGKDNDRGNETFRPG 480
 Db 415 TLOCKIKIINRWQEVGKAMYAPPIEQIRCSNITGLLTRDGGEDTDND-TEIFRPG 473

Qy	481	G	D	M	R	D	N	R	S	E	L	Y	K	Y	K	V	K	I	E	P	L	G	V	A	P	T	K	A	K	R	V	V	Q	R	E	B	A	V	G	L	G	A	M	F	G	F	L	G	A	A	G	S	T	M	G		540		
Db	474	G	D	M	R	D	N	R	S	E	L	Y	K	Y	K	V	T	I	E	P	L	G	V	A	P	T	K	A	K	R	V	V	Q	R	E	K	R	A	-	A	I	G	A	L	F	L	G	F	L	G	A	A	G	S	T	M	G		532
Qy	541	A	A	S	V	T	L	T	V	O	A	R	L	L	S	G	I	V	Q	Q	S	N	L	R	A	T	E	A	O	H	L	L	O	L	T	V	M	G	I	K	O	L	A	R	I	L	A	V	E	R	Y	L	K	O	O		600		
Db	533	A	A	S	V	T	L	T	V	O	A	R	L	L	S	G	I	V	Q	Q	N	N	L	R	A	T	E	A	O	H	L	L	O	L	T	V	M	G	I	K	O	L	A	R	V	I	A	V	E	R	Y	L	K	O	O		592		
Qy	601	L	L	G	W	C	S	G	K	L	C	T	T	T	V	P	N	A	S	W	S	K	N	K	T	L	E	A	I	N	N	N	T	W	M	Q	D	K	E	I	D	N	T	S	L	I	Y	S	L	T	E	S	P		660				
Db	593	L	L	G	W	C	S	G	K	L	C	T	T	T	V	P	N	A	S	W	S	-	N	K	S	-	D	D	I	N	N	N	T	W	M	Q	E	R	E	I	D	N	T	S	L	I	Y	S	L	L	E	K	S		651				
Qy	661	I	O	E	K	N	O	E	L	E	L	D	K	W	A	N	L	N	N	F	D	I	S	N	M	I	Y	I	K	I	F	I	M	I	V	G	G	I	V	A	R	I	V	P	V	L	S	I	N	V	R	Q		720					
Db	652	T	Q	E	K	N	O	E	L	E	L	D	K	W	A	S	I	N	N	F	D	I	T	N	N	I	Y	I	K	I	F	I	M	I	V	G	G	I	V	A	R	I	V	P	A	V	L	S	I	N	V	R	Q		711				
Qy	721	G	Y	S	P	L	S	F	O	T	R	L	P	A	P	C	P	O	R	P	E	E	G	D	R	D	R	S	G	L	L	V	D	G	F	L	T	I	I	W	D	L	S	L	C	F	S	V	H		780								
Db	712	G	Y	S	P	L	S	L	T	R	P	P	V	P	G	P	O	R	P	E	E	G	D	R	D	T	S	G	R	L	V	H	G	F	L	A	I	I	W	D	L	S	L	F	S	Y	H		771										
Qy	781	L	R	D	L	L	I	V	T	R	I	V	E	L	G	R	G	E	W	I	K	Y	M	N	L	L	Q	Y	S	O	E	L	K	N	S	A	V	S	L	F	N	A	T	A	I	A	V	A	E	G	T	D	R	V		840			
Db	772	-	R	D	L	L	I	A	A	R	I	V	E	L	G	R	G	E	W	I	K	Y	M	N	L	L	Q	Y	S	O	E	L	K	S	S	A	V	S	L	N	A	T	A	I	A	V	A	E	G	T	D	R	V		830				
Qy	841	Q	V	L	O	R	V	G	R	A	L	I	H	I	P	T	R	I	O	G	L	E	R	A	L		866																																
Db	831	E	V	L	O	R	A	G	R	A	L	I	H	I	P	T	R	I	O	G	L	E	R	A	L		856																																

RESULT 9

RESOL 3
AAR79173

AAK 79173
ID AAK 79173

XX
ID AAR/91/3 standard; protein

XX
AC
7709173.

AC
AAR/91/3;
VV

XX
DE 00-MAR-100

09-MAR-1999

10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532

DE HIV virus-1 isolate MN gp120 sequence.

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

KW HIV virus-1; gp120; therapeutic; immunogen;

3333

OS Synthetic.

XX

FH Key

FT Protein

FT /note= "HIV-1

FT Misc-difference 301. .400

```
FT /note= "V3 loop"
```

FT Region 401. .800

FT /not

FT	Region	419. .436
FT	Region	419. .436

FT
region
123
/note= "CD4 binding region consensus sequence"

PT	gpl20 peptide.
xx	
PS	Disclosure; Fig 1; 55pp; English.
XX	
CC	Peptide analogues of HIV virus antigen contain a plurality of peptide
CC	monomers, each comprising the gpl20 CD4 binding region consensus sequence
CC	(amino acids 419-436) (see AAR79172). The synthetic peptomers are useful
CC	as therapeutic agents, immunogens in vaccine compositions or for the
CC	diagnosis of disease. Since the analogues maintain the conformation of
CC	the native proteins from which they are derived (the consensus sequence),
CC	they are useful against infection by various HIV-1 and HIV-2 isolates
CC	(see AAR79174-84)
XX	
SQ	Sequence 856 AA;

PT Peptide analogues useful as therapeutic agents, immunogens, or for
PT diagnosis of disease - maintain conformation of native protein, e.g. HIV
PT

fractionating a crude gp120 preparation containing full-length, glycosylated gp120 using ion exchange chromatography so as to provide a first collection of fractions. A fraction from the first collection is selected that exhibits specific binding affinity for CD4 peptide, thereby producing a first fractionated material. The first fractionated material is fractionated by hydrophobic interaction chromatography so as to provide a second collection of fractions from which a second collection is selected that exhibits specific binding affinity for CD4 peptide. This second fraction is fractionated by size exclusion chromatography so as to provide a third collection of fractions exhibiting specific binding affinity for CD4 peptide, thereby providing the purified gp120. The purified gp120 can be used for antibody production and in vaccines. (Updated on 17-OCT-2003 to standardise OS field)

XX SQ Sequence 854 AA;

Query Match 84.8%; Score 3913; DB 2; Length 854;

Best Local Similarity 86.1%; Pred. No. 5.9e-206;

Matches 746; Conservative 35; Mismatches 73; Indels 12; Gaps 8;

QY 1 MRVKGIRRNQHWGCTMLLGLLMICSAATEKLWTVVYGVVPMKEATTLFCASDAKAY 60
DB 1 MRVKGIRRNQHWGCTMLLGLLMICSAATEKLWTVVYGVVPMKEA-TTLFCASDAKAY 59
QY 61 DTEAHNVWATHACVPTDPNPQEVVELNVNTEFNWKNMVEQMHEDIISLWDSLPKVCV 120
DB 60 DTEAHNVWATHACVPTDPNPQEVVELNVNTEFNWKNMVEQMHEDIISLWDSLPKVCV 119
QY 121 LTPCLVTLNCTDLRNTNTNNSDNNSEGITKGGEMKNCNFNATISGDKMKQKEYAL 180
DB 120 LTPCLVTLNCTDLRNTNTNNSDNNSEGITKGGEMKNCNFNATISGDKMKQKEYAL 179
QY 181 LYKLDIEPIDNDNTSVYLSICNTSVITQACPKISFEPIPIHYCAPAGFAILKCNCKKFSG 240
DB 180 LYKLDIVSIDNSTSVYLSICNTSVITQACPKISFEPIPIHYCAPAGFAILKCNCKKFSG 239
QY 241 KGSCKXNVSTVQCTHGRPVVSTOLLNGSLAEBEVIRSENFTNNAKTIIVQLRBPVKIN 300
DB 240 KGSCKXNVSTVQCTHGRPVVSTOLLNGSLAEBEVIRSEFTDNAKTIIVHLNESVQIN 299
QY 301 CSRPNNTKRSIPMGCEAPFTVTCOIGDIROACHNLSKTNWTNALKOVVEKLGEOPNKT 360
DB 300 CTRPNYKRRRIHIGPGRFTYTKNIIGTIRQACHNISRAKNWATLQIVSKLKEQFKNK 359
QY 361 KIYFTSSGSDPEIVTHSFNCAGEFFCYNTQTFDSIWNSENGTWNITRGLNNTGRNDTI 420
DB 360 TIVFNQSSGSDPEIVMHSFNCGE-FYCNTPSFNSTWNGNSTWN-----NITGSNNI 413
QY 421 TLPCKRIKQIINRQWGVKAMYPPIKGNISCSNITGLLTRDGGKDDNDRDGNETFRPG 480
DB 414 TLOCKIKQIINRQWGVKAMYPPIEQIRCSNITGLLTRDGGKDDTND-TEIFRPG 472
QY 481 GGMEDNWRSELKYKVKIEPIGVAPTKAKRRVQREERAVGIGAMFFGLGAGSTMG 540
DB 473 GGMEDNWRSELKYKVKVITIEPIGVAPTKAKRRVQREKKA-AIGALFLGLGAGSTMG 531
QY 541 AASVTLTVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVMGIIQLOARILAVERYLKDDQ 600
DB 532 AASVTLTVQARQLLSGIV-QONLLRAIEAQHMLQLTVMGIIQLOARILAVERYLKDDQ 590
QY 601 LLGIWGCCKGLICTTTVPWNASKNKTLEAIWNNMTWMQDKEDINTYLSLYSLEESP 660
DB 591 LLGIWGCCKGLICTTTVPWNASKS-NKSLDDIWNNTWMQWEREIDNTYLSLYSLEKSQ 649
QY 661 IQEKEQELLEDKQANLWNNFDSINWLYIKIFIMVGLVGLRIVFVVLIVNRVQ 720
DB 650 TQEFKEQELLEDKQASLWNNFDSINWLYIKIFIMVGLVGLRIVFVVLIVNRVQ 709
QY 721 GYSPLSFQTLRPAKPRDPEEIEEGDRDRDRSGLLVDPGLTLIWDLRSLSLFSYHR 780
DB 710 GYSPLSLQTRPPVGRDPEEGIEEGDRDRDTSGLRVHGFLLAIWDLRSLSLFSYH 769
QY 781 LRDLILLIVTRIVELLGRRGWEILKYWNLLQYWSQELKNSAVSLFNATAIAVAGTDRVI 840

DB 770 -RDLIIARIVSELGRRGWEVLYKYNWLLQYWSQELKSSAVSLFNATAIAVAGTDRVI 828
QY 841 QVLQVRGRALLHIPTIRIQGLERALL 866
DB 829 EVLQVRGRALLHIPTIRIQGLERALL 854
RESULT 13
ABR55495
ID ABR55495 standard; protein; 856 AA.
XX ABR55495;
XX 11-AUG-2003 (first entry)
XX Amino acid sequence of a HIV envelope protein.
XX Polymorphic marker; host response; HIV; AIDS related virus; vaccine;
XX hepatitis related virus; HCV; HBV; drug resistance; envelope.
XX Human immunodeficiency virus.
XX WO2003035097-A1.
XX 01-MAY-2003.
XX 23-OCT-2002; 2002WO-AU001450.
XX 23-OCT-2001; 2001AU-00008425.
XX (EPIP-) EPIP PTY LTD.
XX Mallal S;
XX WPI; 2003-449231/42.

Determining the influence of variation in host genes on the selection of microorganisms with protein substitutions, comprises typing individuals of a cohort infected with a microorganism for an intrinsic polymorphic marker.

Claim 22; Page 90-91; 157pp; English.

The specification describes a method of determining the influence of variation in host genes on selection of microorganisms with protein substitutions. The method comprises typing all individuals of a population of patients infected with a microorganism for at least one selected intrinsic polymorphic marker involved in the host response to the presence of the microorganism. The method is useful for examining selective pressures confronting a wide range of organisms that exhibit pathogenic traits in a host, such as bacteria, fungi, mycobacterium, viruses and virus-like particles; for examining microorganisms that have adapted to evolve rapidly, including HIV and AIDS related viruses and the hepatitis related viruses such as HCV and HBV. The method is useful for designing a vaccine to prevent or delay the emergence of drug resistance in patients treated with a particular drug specific for a microorganism where the drug affects the replication of the microorganism at the nucleotide or amino acid level. The present sequence represents a HIV envelope protein, which is expected to provide optimal cytotoxic T lymphocyte (CTL) induced therapeutic protection to the cohort examined in that study

SQ Sequence 856 AA;

Query Match 84.4%; Score 3896; DB 6; Length 856;

Best Local Similarity 84.8%; Pred. No. 5e-205;

Matches 738; Conservative 45; Mismatches 69; Indels 18; Gaps 9;

QY 1 MRVKGIRRNQHW-WCGMTMLLGLLMICSAATEKLWTVVYGVVPMKEATTLFCASDAK 59
DB 1 MRVKGNNQHLWKGWKWGTMLLGLLMICSAATEKLWTVVYGVVPMKEATTLFCASDAK 60

Db 415 GSNSTITLOCKIOLINMMQEVGKAMYAPPIEQIRCSNITGLLTRDGNDDT-NNT 473
 Qy 475 ETRPFGGDMRDNWSELYKYVVKIEPLGVAFTKAKRRVVQREERAVGLGAMFFGFLGA 534
 Db 474 ETRPFGGDMRDNWSELYKYVVKIEPLGVAFTKAKRRVVQREKKA-AIGALFLGFLGA 532
 Qy 535 AGSTMGAASVTITVQARQLLSGIVQOQSNLLRAIEAQHLLQITVWGIKOLQARILAVR 594
 Db 533 AGSTMGAASMTITVQARQLLSGIVQOQSNLLRAIEAQHLLQITVWGIKOLQARILAVR 592
 Qy 595 YLKDQQLLGTWCGSKLICTTTPWNASWKNKTLEAINWMTMOWDKEIDNVTSLIYS 654
 Db 593 YLKDQQLLGTWCGSKLICTTTPWNASWS-NKSQEDIMNMTMOWEREIDNTTSTIYE 651
 Qy 655 LIBESPTQEKNEQLELDKLANLWNFDISNWLWYIKIFIMVGLVGLRIVFVVLISI 714
 Db 652 LLEKSNQEKNEQLELDKLANLWNFDTIWNWYIKIFIMVGLVGLRIVFVVLISI 711
 Qy 715 VNRVQGYSPSLQTRLPAPRGDPREEIEEGGDRDRDTSGLVDFLTIWDLRLSLC 774
 Db 712 VNRVQGYSPSLQTRPPVPGDPREEIEEGGDRDRDTSGLVDFLTIWDLRLSLC 771
 Qy 775 LFSYHRLDLLIVTRIVELLGRGWEILKYWNLLQYNSOELKNSAVSLFNATAVAE 834
 Db 772 LFSYHRLDLLIAARIVELLGRGWEILKYWNLLQYNSOELKNSAVSLFNATAVAE 831
 Qy 835 GTRDVITQVQRGRLIHIPTIRQGLERALL 866
 Db 832 GTRDVIEVQRGRLIHIPTIRQGLERALL 863

RESULT 15
 ADK14406
 ID ADK14406 standard; protein; 847 AA.
 XX
 AC ADK14406;
 XX
 DT 03-JUN-2004 (first entry)
 XX
 DE HIV wild-type gp160 protein.
 XX
 KW HIV; gp120; gp41; glycoprotein; AIDS; gp140; variable loop;
 KW vaccine; viral disease; HIV infection; cytotoxic T lymphocytes;
 KW neutralising antibody.
 XX
 OS Human immunodeficiency virus 1; isolate JR-FL.
 XX
 PN US6710173-B1.
 XX
 PD 23-MAR-2004.
 XX
 PF 23-JUN-2000; 2000US-00602864.
 XX
 PR 25-JUN-1999; 99US-0141168P.
 XX
 PA (PROG-) PROGENICS PHARM INC.
 PA (ADAR-) ADARC AARON DIAMOND AIDS RES CENT.
 XX
 PI Binley JM, Schuelke N, Olson WC, Maddon PJ, Moore JP;
 XX WPI; 2004-256072/24.
 XX
 PT A nucleic acid encoding a modified viral env gene of an HIV-1 isolate,
 PT useful as a vaccine for treating or reducing the severity of a viral
 PT disease, e.g. HIV infection, or reducing the likelihood of a subject
 PT becoming infected with a virus.
 XX
 PS Disclosure; Page; 51pp; English.
 XX
 CC The invention relates to a nucleic acid encoding a modified viral
 CC envelope glycoprotein gene of an HIV-1 isolate comprising a nucleotide
 CC segment that encodes a modified form of an HIV-1 gp120 and gp41 complex,

CC where the modifications comprise a V35C, Y39C, W44C, P484C, G486C, A488C,
 CC F489C, T490C, or A492C mutation in gp120 and a D580C, W587C, T596C,
 CC V599C, P600C, or W601C mutation in gp41, such mutations being numbered by
 CC reference to the gp160 (precursor of gp41 and gp120) HIV-1 isolate JR-FL
 CC and the modifications resulting in a disulphide bond between gp120 and
 CC gp41 which stabilises the otherwise non-covalent gp120-gp41 interaction.
 CC The stabilization results in enhanced binding of the disulphide-bonded
 CC gp120-gp41 complex to HIV-1 neutralising antibodies and reduced binding
 CC of the complex to HIV-1 non-neutralising antibodies. Also included are a
 CC non-replicating viral vector comprising the nucleic acid cited above, a
 CC replicable vector comprising the nucleic acid, a host cell comprising the
 CC modified gp120 is further characterised by the absence of one or more of
 CC the variable loops present in the wild type gp120, these mutants being
 CC further described as gp140, a mutant of gp160 which has the proteolytic
 CC cleavage site between gp120 and gp41 mutated and a stop codon inserted
 CC prior to transmembrane domain of gp41 resulting in a 140kDa protein. The
 CC absent variable loop comprises V1, V2, V3 or its combination, preferably
 CC V1 and V2. The modified gp120 is further characterised by the absence or
 CC presence of one or more canonical glycosylation sites present in wild
 CC type gp120. One or more canonical glycosylation sites are absent from the
 CC V1V2 region of the gp120. The nucleic acids and proteins are useful as
 CC vaccines for treating a viral disease, e.g. HIV infection (AIDS),
 CC reducing the likelihood of a subject becoming infected with a virus,
 CC reducing the severity of a viral disease, or stimulating or enhancing in
 CC a subject production of antibodies or cytotoxic T lymphocytes. The
 CC present sequence is wild-type HIV gp160 protein. NOTE: This sequence is
 CC not shown in the specification but was obtained from genbank reference
 CC U63632 (DNA entry, proteins shown).
 XX
 SQ Sequence 847 AA;

Query Match 83.3%; Score 3845.5; DB 8; Length 847;
 Best Local Similarity 84.1%; Pred. No. 2.9e-202;
 Matches 728; Conservative 54; Mismatches 65; Indels 19; Gaps 6;
 Qy 1 MRVKGIRRNQYHWGWTMLLGLLMICSAATEKLWVTYVYGVVPVWKEATTLFCASDAKAY 60
 Db 1 MRVKGIRKSYQYVLMKGGTLLGLMTCSAVEKLWVTYVYGVVPVWKEATTLFCASDAKAY 60
 Qy 61 DTEAHNVWATHACVPTDNPQVEVLNVTFENFMWKNVVEOMHEDIISLWDSQSLKPCVK 120
 Db 61 DTEAHNVWATHACVPTDNPQVEVLNVTFENFMWKNVVEOMHEDIISLWDSQSLKPCVK 120
 Qy 121 LTPCLVTLNCTDLRNTNTNNDNNNSSEGTIKGEMKNCFSNFIATSTGDKMQKEYAL 180
 Db 121 LTPCLVTLNCKDV-NATNTN-----DSEGTMERGEIKNCFSNFIATSTGDKMQKEYAL 172
 Qy 181 LYKLDIEPDNDNTSYRLISNTSVITQACPKISFEPPIHYCAPAGFAILKCNKDFSG 240
 Db 173 FYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFEPPIHYCAPAGFAILKCNKDFSG 232
 Qy 241 KGSCKNVSTVQCTHGIRPVVSTOLLNGSLAEVEVIRSENFTNNAKTIIVQLREPVKIN 300
 Db 233 KGPCKNVSTVQCTHGIRPVVSTOLLNGSLAEVEVIRSDNFNTNNAKTIIVQLKESVEIN 292
 Qy 301 CSRPNNTNRKSIIPMGPRAFYTTGQIIGDIRQAHNCISKTNTNALKQVVEKLGEOFNKT 360
 Db 293 CTRPNNTNRKSIHIGPRAFYTTGQIIGDIRQAHNCISRAKWNDLTKQIVIKLREQPENK 352
 Qy 361 KIVFTNSSGGDPEIVTHSFNCAGEFFYCNVTQQLFDSIWNSENGTWNITRGLNNTGRDITI 420
 Db 353 TIVFNHSSGGDPEIVMHSFNCGEFFYCNSTQLFNSNTN-----NNTGSENNTEG-TI 405
 Qy 421 TLPCLRIKQIINWQEVGKAWYAPPIKGNISCSNITGLLLTRDGGKDNDSRDNGNEFRPG 480
 Db 406 TLPCLRIKQIINWQEVGKAWYAPPIKQIRCSNITGLLLTRDGGGINEN---GTEIFRPG 462
 Qy 481 GGDMDRDNWRSSELYKYKVVKIIEPLGVAFTKAKRRVVQREERAVGLGAMFFGFLGAAGSTM 540
 Db 463 GGDMDRDNWRSSELYKYKVVKIIEPLGVAFTKAKRRVVQREERAVGLGAMFFGFLGAAGSTM 522
 Qy 541 AASVTITVQARQLLSGIVQOQSNLLRAIEAQHLLQITVWGIKOLQARILAVERYLKDDQ 600

Db	523	AAAMTLTVQARLLLSGIVQQNLLRAIEAQRMQLQTVWGIQLOQARVLAVERYLGDQQ	582
Qy	601	LLGTWGCCKLICTTTVPWNASKXKTLKAIWNNMTMOWDKEIDNYTSLIYSLIEESP	660
Db	583	LLGIWGCCKLICTTAVPWNASWS-NKSLDRIWNNMTMWEWEREIDNYTSEIYTLIEESQ	641
Qy	661	IQOEKNEQELLELDKWANLWNPFDISNWLWYIKIFIMIVGGLVGLRIVFVVLISIVNRVQ	720
Db	642	NQOEKNEQELLELDKWLNNWFDITKWLWYIKIFIMIVGGLVGLRLVFTVLSIVNRVQ	701
Qy	721	GYSPLSFQTLPPAPRGDPREEIEEGGDRDRDRSGLLVDGFLTLIWVDLRSCLFSYHR	780
Db	702	GYSPLSFQTLPPAPRGDPREEIEEGGDRDRDRSGRLVNGFLALIWVDLRSCLFSYHR	761
Qy	781	LRDLLLIVTRIVELLGRGWEIILKYWNLLQYWSQELKNSAVSLFNATAIATAVAEGTDRVI	840
Db	762	LRDLLLIVTRIVELLGRGWEIILKYWNLLQYWSQELKNSAVSLFNATAIATAVAEGTDRII	821
Qy	841	QVLQVRGALLHIPTRIQGLERALL	866
Db	822	EALQTYTRAILHIPTRIQGLERALL	847

Search completed: April 11, 2005, 11:08:59
Job time : 295.737 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2005, 10:57:57 ; Search time 65.6201 Seconds
(without alignments)
1269.790 Million cell updates/sec

Title: US-09-762-261C-1

Perfect score: 4617

Sequence: 1 MRVKGIRRYQHWWGWTML.....GRALLHIPTRIQGLERALL 866

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3977.5	86.1	859	1 VCLJMN	env polyprotein pr
2	3862	83.6	852	2 T12016	env polyprotein pr
3	3853.5	83.5	847	2 S13289	env protein - huma
4	3845.5	83.3	847	2 T09448	env polyprotein pr
5	3839.5	83.2	843	1 H44001	env polyprotein pr
6	3775.5	81.8	861	1 VCLJSC	env polyprotein pr
7	3760	81.4	852	1 VCLJBR	env polyprotein pr
8	3753.5	81.3	861	1 VCLJLV	env polyprotein pr
9	3744	81.1	856	1 VCLJH3	env polyprotein pr
10	3731	80.8	856	1 VCLJVL	env polyprotein pr
11	3716	80.5	854	2 S13288	env polyprotein - huma
12	3710.5	80.4	851	2 S33985	env polyprotein -
13	3708	80.3	868	1 VCLJH4	env polyprotein pr
14	3703.5	80.2	861	1 VCLJKB	env polyprotein pr
15	3700	80.1	856	1 VCLJ3W	env polyprotein pr
16	3688.5	79.9	855	1 VCLJ42	env polyprotein pr
17	3523.5	76.3	859	2 T01672	env polyprotein pr
18	3500.5	75.8	853	2 S54384	env polyprotein pr
19	3500.5	75.8	855	1 VCLJZR	env polyprotein pr
20	3398	73.6	856	1 A44963	env polyprotein pr
21	3396	73.6	846	1 VCLJND	env polyprotein pr
22	3117	67.5	729	1 VCLJXK	env polyprotein pr
23	2746	59.5	854	1 VCLJ91	env polyprotein pr
24	2300.5	49.8	877	2 S49197	env polyprotein pr
25	2198	47.6	863	2 A50314	env polyprotein pr
26	2179	47.2	506	2 A40218	env polyprotein pr
27	1987.5	43.0	495	2 S31493	env polyprotein pr
28	1963	42.5	443	2 C41621	env polyprotein pr
29	1955	42.3	445	2 A41621	env polyprotein pr

RESULT 1

VCLJMN

env polyprotein precursor - human immunodeficiency virus type 1 (isolate MN)

N;Alternate names: coat polyprotein

C;Species: human immunodeficiency virus type 1, HIV-1

A;Note: host Homo sapiens (man)

C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Feb-1997

C;Accession: A28922

R;Gurgo, C.; Guo, H.G.; Franchini, G.; Aldovini, A.; Collalti, B.; Farrell, K.; Wong-Stat

Virology 164, 531-536, 1988

A;Title: Envelope sequences of two new United States HIV-1 isolates.

A;Reference number: A28922; MUID:88219542; PMID:3369091

A;Accession: A28922

A;Molecule type: DNA

A;Residues: 1-859 <GUR>

C;Genetics:

A;Gene: env

C;Superfamily: type E retrovirus env polyprotein

C;Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane protei

F;1-29/Domain: signal sequence #status predicted <SIG>

F;30-859/Product: env polyprotein #status predicted <EP>

F;87,129,135,140,141,146,165,191,202,246,267,281,294,300,336,343,359,365,389,395,401,

Query Match 86.1%; Score 3977.5; DB 1; Length 859;

Best Local Similarity 87.0%; Pred. No. 4.2e-277;

Matches 753; Conservative 36; Mismatches 70; Indels 7; Gaps 6;

QY 1 MRVKGIRRYQHWWGWTMLGLLMICSAATEKLWVTYVYGVVPVWKEATTLFCASDAKAY 60

Db 1 MRVKGIRRYQHWWGWTMLGLLMICSAATEKLWVTYVYGVVPVWKEATTLFCASDAKAY 60

QY 61 DTEAHNVWATHACVPTDPNPQVELNVVTENFMNKNMVEQMHEDIISLWDSLRPCVK 120

Db 61 DTEAHNVWATHACVPTDPNPQVELNVVTENFMNKNMVEQMHEDIISLWDSLRPCVK 120

QY 121 LTPLCVTLNCTDLRNTTNNSTNNNSSEGTIKGEMKNCNFNATSIATSGDKMKEAYAL 180

Db 121 LTPLCVTLNCTDLRNTTNNSTNNNSSEGTIKGEMKNCNFNATSIATSGDKMKEAYAL 180

QY 181 LYKLDIEPTDNNTSVRLISCVTSVITQACPKISFPIPIHYCAPAGFAILKCNCKKFSG 240

Db 181 LYKLDIVSDNDSTSVRLISCVTSVITQACPKISFPIPIHYCAPAGFAILKCNCKKFSG 240

QY 241 KGSCKNVSTVQCTHGRPVVSTOLLNGSLAEVVRSENFTNNAKTIIVQLREPVKIN 300

Db 241 KGSCKNVSTVQCTHGRPVVSTOLLNGSLAEVVRSENFTNNAKTIIVHLNESVQIN 300

QY 301 CSRPNNTKRSIPMGFGRAFYTTGIIIGIRQAHNCISKTNTWTNALKQVVEKGEQFNKT 360

Db 301 CTRPNYKRRKHIGFGRAFYTTKNIIGIRQAHNCISRAKNDTLRQIVSKLKEQFNK 360

QY 361 KIVFTNSSGGDPEIVTHSFNCAGEFFYCNTTQLFDSIWNSENGTWNITRGLNNTGRNDTI 420

```
Db 361 TIVFNSSGGDPEIVHVSFNCGGEFFYCNTPSPFNSTWNG--NNTW--TKGNNTTSGNNNI 417
Qy 421 TLPCRILQIILNRQWGVCKANYAPIKGNISCSNITGLLLTRDGGKDDNSRDGNEFRPG 480
Db 418 TLOCKIQLIINMQWGVCKANYAPIEIGQIRCSNITGLLLTRDGGKDDTND--TEIFRPG 476
Qy 481 GGDMDNRWSELYKYKVVKIEPLGVAPTAKRRVVQREERAAGLGMFFGLGAAGSTWG 540
Db 477 GGDMDNRWSELYKYKVVVTEPLGVAPTAKRRVVQREKEA--AIGALFLGLGAAGSTWG 535
Qy 541 AASVTLTVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVWGIKQLOARILAVRYLKQQ 600
Db 536 AASVTLTVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVWGIKQLOARILAVRYLKQQ 595
Qy 601 LLGIWCSGKLICTTTPVWNASKNKTLKAIWNNMTWQDKEIDNYSLIYSLIEEP 660
Db 596 LLGIWCSGKLICTTTPVWNASWS--NKSLEDDNNMTWQDKEIDNYSLIYSLIEKSQ 654
Qy 661 IQQEKNEQELLELDKWNLMNWFDSINLWYIKIFIMIVGGLVGLRIVFVLSIVNRVQ 720
Db 655 TQQEKNEQELLELDKWNLMNWFDSINLWYIKIFIMIVGGLVGLRIVFVLSIVNRVQ 714
Qy 721 GYSPLSFQTRLPAPRGPDRPEIEEGGDRDRDRSGLLVGDFUTLWDLRSICLSYHR 780
Db 715 GYSPLSFQTRLPAPRGPDRPEIEEGGDRDRDRSGLLVGDFUTLWDLRSICLSYHR 774
Qy 781 LRDLILIVTRIVELGRRGWEILKYWNLLQYWSQELKNSAVSLFNATAVAEGTDRVI 840
Db 775 -RDLILIAIVELGRRGWEILKYWNLLQYWSQELKNSAVSLFNATAVAEGTDRVI 833
Qy 841 QVLQVRGALLHIPTRIQGLERALL 866
Db 834 EVLQVRGALLHIPTRIQGLERALL 859

RESULT 2
T12016
envelope glycoprotein - human immunodeficiency virus type 1 (strain sc14.3)
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T12016
R:McCutchan, F.E.; Sanders-Buell, E.; Salminen, M.O.; Carr, J.K.; Sheppard, W.H.
AIDS Res. Hum. Retroviruses 14, 329-337, 1998
A:Title: Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in S
A:Reference number: Z17379; MUID:98178716; PMID:9519894
A:Accession: T12016
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-852 <MCC>
A:Cross-references: UNIPROT:O41883; EMBL:U90934; NID:g2351783; PIDN:AAC59271.1; PID:g235
C:Genetics:
C:Gene: env
C:Superfamily: type B retrovirus env polyprotein

Query Match 83.6%; Score 3862; DB 2; Length 852;
Best Local Similarity 84.0%; Pred. No. 8e-269;
Matches 728; Conservative 54; Mismatches 69; Indels 16; Gaps 6;
Qy 1 MRVKGIRRNQHWGWTMLGLMLTCSATEKLWTVTVYGVVPWKEATTLFCASDAKAY 60
Db 1 MRVKGIRKNCQHLWKWGTMLGLMLTCSAABQLRVTVYGVVPWKEATTLFCASDAKAY 60
Qy 61 DTEAHNVWATHACVPTDPPQVELNVVTENFNWKNMVEQMHEDIISLWDSQSLKPCVK 120
Db 61 DTEAHNVWATHACVPTDPPQVELNVVTENFNWKNMVEQMHEDIISLWDSQSLKPCVK 120
Qy 121 LTPLCVTLNCTD--LRNTTNTNSDNNNSSEGTIRKGEKNCSEFNATISIGDKMOKEYA 179
Db 121 LTPLCVTLNCTDYLNDTNT-----TSNNGGMEGGEIKNCSEFNATIRGNKVQKEYA 173
Qy 180 LLYKLDIEPINDNTSYRLISCNSTSVITQCPKISPEPIHIHYCTPAGFALLCKDCKFN 239
```

```
Db 174 LFYKLDVDPIDNTTTSYRLINCNTSVITQCPKVSFPIPIHYCTPAGFALLCKDCKFN 233
Qy 240 GKSGCKNVSTVQCTHGIRPVVSTQLLNLGSLAEVEVVIRSENFNTNNAKTIIVQLRPPVKI 299
Db 234 GTGFCNTNVSTVQCTHGIRPVVSTQLLNLGSLAEVEVVIRSENFNTNNAKTIIVQLNESVEI 293
Qy 300 NCSRPNNTNRKSIIPMGPGRAFYTTGQIIGDIRQAHCHNISKTNNWTNALQVVEKLGQFNK 359
Db 294 NCTRPNNNTNRKSIHIGPGQALYATGAIIGDIRQAHCHNISRAKWNNTLKQIVKLVQVQFN 353
Qy 360 TKLIVFTNSGGDPEIVTHSFNCAGFFYCNTPQLPDSIWNSENGTWNITRGLNNTGRNDT 419
Db 354 KTIIFNQSSGGDPEIVHVSFNCGGEFFYCNTPKLFNSTW--MFNNTWNTD---EDTEGNGT 409
Qy 420 ITPLCRIKQIINRWQBVGMKAMVAPPIKGNISCSNITGLLLTRDGGKDDNSRDGNETFRP 479
Db 410 ITPLCRIKQIINRWQBVGMKAMVAPPIRQIRCSNITGILLTRDGGTNNST---NETFRP 466
Qy 480 GGDMDNRWSELYKYKVVKIEPLGVAPTAKRRVVQREERAAGLGMFFGLGAAGSTM 539
Db 467 GGDMDNRWSELYKYKVVKIEPLGVAPTAKRRVVQREKRAVGIGALFLGLGAAGSTM 526
Qy 540 GAASVTLTVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVWGIKQLOARILAVRYLKQQ 599
Db 527 GAASVTLTVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVWGIKQLOARILAVRYLKQQ 586
Qy 600 QLLGIWCSGKLICTTTPVWNASWSKNKTLKAIWNNMTWQDKEIDNTSYSLIYSLIEES 659
Db 587 QLLGIWCSGKLICTTTPVWNASWS--NKSLEDKIWNNTWMEWEREIDNTSYSLIYSLIEES 645
Qy 660 PIQEKNEQELLELDKWNLMNWFDSINLWYIKIFIMIVGGLVGLRIVFVLSIVNRVR 719
Db 646 QNQQVKNQELLELDKWNLSWGFDTKWLWYIKIFIMIVGGLVGLRIVFVLSIVNRVR 705
Qy 720 QYSPLSFQTRLPAPRGPDRPEIEEGGDRDRDRSGLLVGDFUTLWDLRSICLSFSYH 779
Db 706 QYSPLSFQTRLPAPRGPDRPEIEEGGDRDRSGLLVGDFUTLWDLRSICLSFSYH 765
Qy 780 RLRLDILLIVTRIVELGRRGWEILKYWNLLQYWSQELKNSAVSLFNATAVAEGTDRV 839
Db 766 RLRLDILLIVARIVELGREGWEALKYWNLLQYWSQELKNSAVSLFNATAVAEGTDRV 825
Qy 840 IQVLQVRGALLHIPTRIQGLERALL 866
Db 826 IEVQRTCRAILHIPTRIQGLERALL 852

RESULT 3
S13289
env protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S13289
R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A.
Nature 348, 69-73, 1990
A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120 c
A:Reference number: S13288; MUID:91043044; PMID:2172833
A:Accession: S13289
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-847 <OBR>
A:Cross-references: UNIPROT:Q75760
C:Superfamily: type E retrovirus env polyprotein

Query Match 83.5%; Score 3853.5; DB 2; Length 847;
Best Local Similarity 84.2%; Pred. No. 3.2e-268;
Matches 729; Conservative 52; Mismatches 66; Indels 19; Gaps 6;
Qy 1 MRVKGIRRNQHWGWTMLGLMLTCSATEKLWTVTVYGVVPWKEATTLFCASDAKAY 60
Db 1 MRVKGIRKSYQVILWKWGTMLGLMLTCSATEKLWTVTVYGVVPWKEATTLFCASDAKAY 60
Qy 61 DTEAHNVWATHACVPTDPPQVELNVVTENFNWKNMVEQMHEDIISLWDSQSLKPCVK 120
```

```
Db 61 DTEHNVWATHACVPTDPNPQEVLENVTFNFMKNMVMVQOEIIISLWDSQKPCVK 120
Qy 121 LTPCLVTLNCTDLRNTNTNNTSSEGIKKGEMKNCFSNIATSIIGDMQKQYKAL 180
Db 121 LTPCLVTLNCKOV-NATNTN-----DSBGMTGEIKNCFSNITTSIRDEVQKAY 172
Qy 181 LYKLDIEPIDNDNTSVRLISCNSTSVITQACPKISFPIPIHYCAPAGFAILKNDKFFSG 240
Db 173 FYKLDVVPIDNNNTSVRLISCDTSVITQACPKISFPIPIHYCAPAGFAILKNDKTFNG 232
Qy 241 KGSCKNVSTVQCTHGRPVVSTQLLNGSLABEEVIRSNFTNNAKTIIVQLREPVKIN 300
Db 233 KGPCKNVSTVQCTHGRPVVSTQLLNGSLABEEVIRSNFTNNAKTIIVQLKESVEIN 292
Qy 301 CSRPNNTKRSIPMGPRAPYTTGQIIGDIROAHCHNISKTNTWNAKQVVEKLGEOFNKT 360
Db 293 CTEPNNTKRSIHIGPRAPYTTGQIIGDIROAHCHNISRAKWDNLKQIVIKLREQFENK 352
Qy 361 KIVFTNSSGDDPEIVTHSFNCAGFPYCNTOFDSIWNSENGTWNITRGLNNTGRNDTI 420
Db 353 TIVFNHSSGDDPEIVHVSFNCAGFPYCNTOFDSIWNSENGTWNITRGLNNTGRNDTI 405
Qy 421 TLPCKRIKQIINRWQEVGKAMYAPPIKGNISCSNITGLLLTRDGGKDDNSRDCNETFRPG 480
Db 406 TLPCKRIKQIINRWQEVGKAMYAPPIKGNISCSNITGLLLTRDGGKDDNSRDCNETFRPG 462
Qy 481 GGDMDRNWSELYKYKVVKIEPLGVAPTAKRRVQREKRAVIGAVFLGFLGAAGSTMG 540
Db 463 GGDMDRNWSELYKYKVVKIEPLGVAPTAKRRVQREKRAVIGAVFLGFLGAAGSTMG 522
Qy 541 AASVTLTVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVMGIKQLOARILAVERYLKDQ 600
Db 523 AASMTLTVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVMGIKQLOARILAVERYLKDQ 582
Qy 601 LLGIWCSGKGLICTTTPVWNASKNKTLEAIWNNTWQMDKEIDNYTSLIYSLIEESP 660
Db 583 LLGIWCSGKGLICTTTPVWNASKNKTLEAIWNNTWQMDKEIDNYTSLIYSLIEESP 641
Qy 661 IQOEKNEQELLELDKQANLWNNFDSINWLYKIFIMIVGGLVGLRIVFVLSIVNRVQ 720
Db 702 GYSPLSFQTLPPAPGRDPEEIEEGDRDRDRSGRLVNGFLALIVWDLRSICLFSYHR 761
Qy 781 LRDLIIIVTRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATAVAEGTDRVI 840
Db 762 LRDLIIIVTRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATAVAEGTDRVI 821
Qy 841 QVLQVRGRALLHIPTRIQGLERALL 866
Db 822 EALQRTYRALLHIPTRIQGLERALL 847
```

RESULT 4

```
T09448
envelope glycoprotein - human immunodeficiency virus type 1 (strain JRFL)
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T09448
R:Pang, S.; Vinters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namazie,
submitted to the EMBL Data Library, July 1996
A:Reference number: Z16673
A:Accession: T09448
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-847 <PAN>
A:Cross-references: UNIPROT:Q75760; EMBL:U63632; NID:gl465777; PID:gl465781
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
```

```
Query Match 83.3%; Score 3845.5; DB 2; Length 847;
Best Local Similarity 84.1%; Pred. No. 1.2e-267;
Matches 728; Conservative 54; Mismatches 65; Indels 19; Gaps 6;
Qy 1 MRVKGIRRNQHWGWTMLLGLLMICSAATEKLWTVTVYGVVWKEATTTLFCASDAKAY 60
Db 1 MRVKGIRKSYQYLVKGGTLLGLLMICSAVEKLWTVTVYGVVWKEATTTLFCASDAKAY 60
Qy 61 DTEAHNVWATHACVPTDPNPQEVLENVTFNFMKNMVMVQOEIIISLWDSQKPCVK 120
Db 61 DTEAHNVWATHACVPTDPNPQEVLENVTFNFMKNMVMVQOEIIISLWDSQKPCVK 120
Qy 121 LTPCLVTLNCTDLRNTNTNNTSSEGIKKGEMKNCFSNIATSIIGDMQKQYKAL 180
Db 121 LTPCLVTLNCKOV-NATNTN-----DSBGMTGEIKNCFSNITTSIRDEVQKAY 172
Qy 181 LYKLDIEPIDNDNTSVRLISCNSTSVITQACPKISFPIPIHYCAPAGFAILKNDKFFSG 240
Db 173 FYKLDVVPIDNNNTSVRLISCDTSVITQACPKISFPIPIHYCAPAGFAILKNDKTFNG 232
Qy 241 KGSCKNVSTVQCTHGRPVVSTQLLNGSLABEEVIRSNFTNNAKTIIVQLREPVKIN 300
Db 233 KGPCKNVSTVQCTHGRPVVSTQLLNGSLABEEVIRSNFTNNAKTIIVQLKESVEIN 292
Qy 301 CSRPNNTKRSIPMGPRAPYTTGQIIGDIROAHCHNISKTNTWNAKQVVEKLGEOFNKT 360
Db 293 CTEPNNTKRSIHIGPRAPYTTGQIIGDIROAHCHNISRAKWDNLKQIVIKLREQFENK 352
Qy 361 KIVFTNSSGDDPEIVTHSFNCAGFPYCNTOFDSIWNSENGTWNITRGLNNTGRNDTI 420
Db 353 TIVFNHSSGDDPEIVHVSFNCAGFPYCNTOFDSIWNSENGTWNITRGLNNTGRNDTI 405
Qy 421 TLPCKRIKQIINRWQEVGKAMYAPPIKGNISCSNITGLLLTRDGGKDDNSRDCNETFRPG 480
Db 406 TLPCKRIKQIINRWQEVGKAMYAPPIKGNISCSNITGLLLTRDGGKDDNSRDCNETFRPG 462
Qy 481 GGDMDRNWSELYKYKVVKIEPLGVAPTAKRRVQREKRAVIGAVFLGFLGAAGSTMG 540
Db 463 GGDMDRNWSELYKYKVVKIEPLGVAPTAKRRVQREKRAVIGAVFLGFLGAAGSTMG 522
Qy 541 AASVTLTVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVMGIKQLOARILAVERYLKDQ 600
Db 523 AASMTLTVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVMGIKQLOARILAVERYLKDQ 582
Qy 601 LLGIWCSGKGLICTTTPVWNASKNKTLEAIWNNTWQMDKEIDNYTSLIYSLIEESP 660
Db 583 LLGIWCSGKGLICTTTPVWNASKNKTLEAIWNNTWQMDKEIDNYTSLIYSLIEESP 641
Qy 661 IQOEKNEQELLELDKQANLWNNFDSINWLYKIFIMIVGGLVGLRIVFVLSIVNRVQ 720
Db 642 NOOEKNEQELLELDKQANLWNNFDSINWLYKIFIMIVGGLVGLRIVFVLSIVNRVQ 701
Qy 721 GYSPLSFQTLPPAPGRDPEEIEEGDRDRDRSGRLVNGFLALIVWDLRSICLFSYHR 780
Db 702 GYSPLSFQTLPPAPGRDPEEIEEGDRDRDRSGRLVNGFLALIVWDLRSICLFSYHR 761
Qy 781 LRDLIIIVTRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATAVAEGTDRVI 840
Db 762 LRDLIIIVTRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATAVAEGTDRVI 821
Qy 841 QVLQVRGRALLHIPTRIQGLERALL 866
Db 822 EALQRTYRALLHIPTRIQGLERALL 847
```

RESULT 5

```
H44001
env polyprotein precursor - human immunodeficiency virus type 1 (strain YU-2)
N:Alternate names: coat polyprotein
N:Contains: coat protein gp120; coat protein gp41
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
```

C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: H44001
R;Li, Y.; Hui, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.
J. Virol. 66, 6587-6600, 1992
A;Title: Complete nucleotide sequence, genome organization, and biological properties of
A;Reference number: A44001; MUID:93021387; PMID:1404605
A;Accession: H44001
A;Molecule type: DNA
A;Residues: 1-843 <LTY>
A;Cross-references: UNIPROT:P35961; GB:M33258
C;Genetics:
C;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F;1-29/DNA: signal sequence #status predicted <SIG>
F;19-35/Region: hydrophobic
F;30-489/Product: coat protein gp120 #status predicted <GP1>
F;490-843/Product: coat protein gp41 #status predicted <GP2>
F;499-515/Region: hydrophobic
F;673-689/Region: hydrophobic
F;738-755/DNA: transmembrane #status predicted <TMN>
F;87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435
Query Match 83.2%; Score 3839.5; DB 1; Length 843;
Best Local Similarity 84.0%; Pred. No. 3.2e-267;
Matches 728; Conservative 46; Mismatches 68; Indels 25; Gaps 7;
QY 1 MRVKGIRRYQHWWGWTMLGLLMICSAATEKLWVTYVYGVVWKEATTLFCASDAKAY 60
DB 1 MRATEIKNYQHLWKGTTGLGLMLCSAAEQWLVTYVYGVVWKEATTLFCASDAKAY 60
QY 61 DTEAHNWAHACVPTDNPQVEVLNVTFNFMNKNMVEQMHEDIISLWDSLKPCVK 120
DB 61 DTEAHNWAHACVPTDNPQVEVLNVTFNFMNKNMVEQMHEDIISLWDSLKPCVK 120
QY 121 LTPCLVTLNCTDLRNTNTNNSTNNSEGTIKGEMKNCFNATISIGDKMQKEYAL 180
DB 121 LTPCLVTLNCTDLRNTNTNNSTNNSEGTIKGEMKNCFNATISIGDKMQKEYAL 173
QY 181 LYKLDIEPIDNDNTSYRLISCVTSVITQACPKISPEPIPIHYCAPAGFAILKCNCKKFSG 240
DB 174 FYNLDVUPI--DNASVRLISCVTSVITQACPKVSPFPIPIHYCAPAGFAILKCNCKKFSG 231
QY 241 KGSKNVSTVQCTHGIRPVVSTQLLNGSLAEVEVIRSENFTNNAKTIIVOLREPVKIN 300
DB 232 TGPCTNVSTVQCTHGIRPVVSTQLLNGSLAEVEVIRSENFTNNAKTIIVOLNESVWIN 291
QY 301 CSRPNNTKRSIPWGPGRFYTTGQIIGDIRQACHNISKNTNNAKTIIVOLREPVKIN 360
DB 292 CTRPNNTKRSIPWGPGRFYTTGQIIGDIRQACHNISKNTNNAKTIIVOLREPVKIN 351
QY 361 K-IVFTNSSGGDPEIVTHSFNCAGEFFYCNVTQTFDSIWNSENGTWNITRGLNNTGRND 419
DB 352 KTIIFNPSSGGDPEIVTHSFNCAGEFFYCNVTQTFDSIWNSENGTWNITRGLNNTGRN 400
QY 420 ITLPCKRIKQIINRWQEVGKAMYPPIKGNISCSNITGLLTDGKDNRSRDNFTFRP 479
DB 401 ITLPCKRIKQIINRWQEVGKAMYPPIRGQIRGTCSSNITGLLTDGKDNFTFRP 457
QY 480 GGGDMRDNRSSELYKYKVKVKEPLGVAPTAKERVVQREERAVGLGALFGFAGAGSTM 539
DB 458 GGGDMRDNRSSELYKYKVKVKEPLGVAPTAKERVVQREERAVGLGALFGFAGAGSTM 517
QY 540 GAASVLTITVOARQLLSGIVQOQSNLLRAIEAQHLLQLTWGIGKQIARLAVERYLKQ 599
DB 518 GAASVLTITVOARQLLSGIVQOQSNLLRAIEAQHLLQLTWGIGKQIARLAVERYLKQ 577
QY 600 QLLGIWCSGKLICTTTVPWNASNKTLLEAIWNNTWQMDKEIDNTSYLSIIES 659
DB 578 QLLGIWCSGKLICTTTVPWNTSWS--NKSINEIWNNTWQMDKEIDNTSYLSIIES 636
QY 660 PIQEQNEBELLELDKWNINWDFITKMLWYIKFIMIVGGILGRIVFVLSIVNRV 719

DB 637 QNOQEKNEQELLALDKWASLWNNFDTKMLWYIKFIMIVGGILGRIVFVLSIVNRV 696
QY 720 QGVSPSLSFOTRLPAPGPDPPEIEBEGGDRDRSGLVDGFLTLIWDLSICLPSYH 779
DB 697 QGVSPSLSFOTRLPAPGPDPPEIEBEGGDRDRSGLVDGFLTLIWDLSICLPSYH 756
QY 780 RLFDLLIVTRIVELLGRGWELKYNWNLQYWSOELKNSAVSLFNATAIAVAEGTDRV 839
DB 757 RLFDLLIVTRIVELLGRGWELKYNWNLQYWSOELKNSAVSLFNATAIAVAEGTDRV 816
QY 840 IQVLQVRGRALLHIPRIQGLERALL 866
DB 817 IEILQRAFRVAVLHIPRIQGLERALL 843
RESULT 6
VCLJSC
env polyprotein precursor - human immunodeficiency virus type 1 (isolate SC)
N;Alternate names: coat polyprotein
C;Species: human immunodeficiency virus type 1, HIV-1
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Feb-1997
C;Accession: B28922
R;Gurgo, C.; Guo, H.G.; Franchini, G.; Aldovini, A.; Collalti, E.; Farrell, K.; Wong-Stat
Virology 164, 531-536, 1988
A;Title: Envelope sequences of two new United States HIV-1 isolates.
A;Reference number: A28922; MUID:88219542; PMID:3369091
A;Accession: B28922
A;Molecule type: DNA
A;Residues: 1-861 <GUR>
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane protei
F;1-29/DNA: signal sequence #status predicted <SIG>
F;30-861/Product: env polyprotein #status predicted <EPP>
F;87,129,135,140,143,159,163,187,198,234,241,262,276,295,301,302,333,340,356,362,386,396,
Query Match 81.8%; Score 3775.5; DB 1; Length 861;
Best Local Similarity 82.3%; Pred. No. 1.3e-262;
Matches 716; Conservative 56; Mismatches 83; Indels 15; Gaps 9;
QY 1 MRVKGIRRYQHWWGWTMLGLLMICSAATEKLWVTYVYGVVWKEATTLFCASDAKAY 60
DB 1 MRVKGIRRYQHWWGWTMLGLLMICSAATEKLWVTYVYGVVWKEATTLFCASDAKAY 60
QY 61 DTEAHNWAHACVPTDNPQVEVLNVTFNFMNKNMVEQMHEDIISLWDSLKPCVK 120
DB 61 DTEAHNWAHACVPTDNPQVEVLNVTFNFMNKNMVEQMHEDIISLWDSLKPCVK 120
QY 121 LTPCLVTLNCTDLRNTNTNNSTNNSEGTIKGEMKNCFNATISIGDKMQKEYAL 180
DB 121 LTPCLVTLNCTDLRNTNTNNSTNNSEGTIKGEMKNCFNATISIGDKMQKEYAL 176
QY 181 LYKLDIEPIDNDNTSYRLISCVTSVITQACPKISPEPIPIHYCAPAGFAILKCNCKKFSG 240
DB 177 FYKLDVVPIDNDNTSYRLISCVTSVITQACPKVSPFPIPIHYCA-RWFAILKCNCKKFSG 235
QY 241 KGSKNVSTVQCTHGIRPVVSTQLLNGSLAEVEVIRSENFTNNAKTIIVOLREPVKIN 300
DB 236 TGPCTNVSTVQCTHGIRPVVSTQLLNGSLAEVEVIRSENFTNNAKTIIVOLREPVKIN 295
QY 301 CSRPNNTKRSIPWGPGRFYTTGQIIGDIRQACHNISKNTNNAKTIIVOLREPVKIN 358
DB 296 CTRPNNTKRSIPWGPGRFYTTGQIIGDIRQACHNISKNTNNAKTIIVOLREPVKIN 355
QY 359 KTKIIVFTNSSGGDPEIVTHSFNCAGEFFYCNVTQTFDSIWNSENGTWNITRGLNNTGRND 418
DB 356 NKTIIIFNRSSGGDPEIVTHSFNCAGEFFYCNVTQTFDSIWNSENGTWNITRGLNNTGRND 411
QY 419 TITLPCRIKQIINRWQEVGKAMYPPIKGNISCSNITGLLTDGKDNFTFRP 476
DB 412 TITLPCRIKQIINRWQEVGKAMYPPIKGVKCSNITGLLTDGKDNFTFRP 471

```
QY 477 FRPGGDMRNRSELYKYKVKVIEPLGVAPTAKKRVVQREBRVAVG-IGAMFFGLGAA 535
Db 472 FRPGGDMRNRSELYKYKVKVIEPLGVAPTAKKRVVQREBRVAVGTTIGAMFFGLGAA 531
QY 536 GSTMGASVTLTVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVWGIKQOARILAVERY 595
Db 532 GSTMGATSMFLTVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVWGIKQOARILAVERY 591
QY 596 LKQDQLLGWCSGKLICTTTVPWNASWSKNTLEAIWNMTMOWDKEIDNTSYLSIYL 655
Db 592 LRQDQLLGWCSGKLICTTTVPWNTSWS-NKSLKDINGNMTWMEWEREIDNTSYLYTL 650
QY 656 IEPSPTQOEKNEQELLELDKWANLWNPDISNWLWYIKIFIMIVGVLGRLRVFVLSIV 715
Db 651 IEBSQOEKNEQELLELDKWANLWNPDISNWLWYIKIFIMIVGVLGRLRVFVLSIV 710
QY 716 NRVQGYSPLSFOTRLPAPGDPDRPEIEBEGDRDRSGLLAVDGLTLIWDLSLCL 775
Db 711 NRVQGYSPLSFOTRLPAPGDPDRPEIEBEGDRDRSGLLAVDGLTLIWDLSLCL 770
QY 776 FSHRLRDLIIIVTRIVELGRGWEILKYWNLLQVWSOELKNSAVSLFNATAIAVAVG 835
Db 771 FSHRLRDLIIIVTRIVELGRGWEILKYWNLLQVWSOELKNSAVSLFNATAIAVAVG 830
QY 836 TDRVIQVQRVGRALLHIPTIRIQLERAL 865
Db 831 TDRVIELLQRAFRAILHIPTIRIQLERAL 860

RESULT 7
VCLJBR
env polyprotein - human immunodeficiency virus type 1 (isolate BR)
N:Alternate names: coat polyprotein
N:Contains: coat protein gp120; coat protein gp41
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 09-Jul-2004
C:Accession: A31667
R:Anand, R.; Thayer, R.; Srinivasan, A.; Nayyar, S.; Gardner, M.; Luciw, P.; Dandekar, S.
Virology 168, 79-89, 1989
A:Title: Biological and molecular characterization of human immunodeficiency virus (HIV-1)
A:Reference number: A94389; MUID:89085613; PMID:2789516
A:Accession: A31667
A:Molecule type: DNA
A:Residues: 1-852 <ANA>
A:Cross-references: UNIPROT:P12488
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: capsid protein; coat protein; glycoprotein; transmembrane protein
F:1-516/Product: coat protein gp120 #status predicted <CP1>
F:517-852/Product: coat protein gp41 #status predicted <CP2>

Query Match 81.4%; Score 3760; DB 1; Length 852;
Best Local Similarity 82.1%; Pred. No. 1.7e-261;
Matches 713; Conservative 54; Mismatches 81; Indels 20; Gaps 8;

QY 1 MRVKGIRRNQHWGW-GTMLLGLMICSATEKLVTVVYGVVPWKEATTLFLCASDAKA 59
Db 1 MRVKGIRKNQVHLRWGCGMMLGLMICSATDKLVTVVYGVVPWKEATTLFLCASDAKA 60

QY 60 YDTEAHNVWATHACVTPDPNPQVELVNVTFENFMWKNVQMHEDIISLWQSLKPCV 119
Db 61 YDTEIHNWATHACVTPDPNPQELVMGNVTENFMWKNVQMHEDIISLWQSLKPCV 120

QY 120 KLTPLCVTLNCTDLRNTVTNTNNDNNSECTIKGEMKNCSFNATISIGDKMOKEYA 179
Db 121 KLTPLCVTLNCHDF----NATNAT----SNSGKMEGEMKNCSFNITTSIRDKMOKEYA 172

QY 180 LLYKLDIEPIND-NTSYRLISNTSVITQACPKISFEPIPIHYCAPAGFALLKNDKK 237
Db 173 LFYKLDIVPIDNKTWTRYRLISNTSVITQACPKYTFEPIPIHYCAPAGFALLKCNKK 232

QY 238 FSGKSGCKNVSTVQCTHGIRPVVSTQLLNGSLAEBEWIRSENFNTNNAKTIIVQLREP 297
```

```
Db 233 FNGTGCCTVSTVQCTHGIRPVVSTQLLNGSLAEBEWIRSENFNTNNAKTIIVQLNESV 292
QY 298 KINCSPPNNNTKRSIPMGPCRAFYYTQIGDIROAHFCNISKTNNTWNAKQVVEKLGEOF 357
Db 293 EINCTPNNNTKRIITMGPGRVVYTTGQIIGDIRRAHCNLSRSKWENTLKQIVTKLRVQF 352
QY 358 NKTKIVFTSSGDDPPIVTHSFNCAGEFFYNTTQLPFDSIWNSENGTWTNITRGLNWTGRN 417
Db 353 KNTKIVFNSSGDDPPIVTHSFNCAGEFFYNTTQLPFDSIWNSENGTWTNITRGLNWTGRN 405
QY 418 DTTPLPCRKIOIINRQEVGKAWYAPPIKGNITSCSNITGLLLTRDGGKDDNRDGNETF 477
Db 406 SPITPLPCRKIOIINRQEVGKAWYAPPIRQIQIKCSNITGLLLTRDGGKNNETD-TEIF 464
QY 478 RFGGDMRNRSELYKYKVKVIEPLGVAPTAKKRVVQREBRVAVGLGAMFFGFLGAAGS 537
Db 465 RFGGDMRNRSELYKYKVKVIEPLGVAPTAKKRVVQREBRVAVGLGAMFFGFLGAAGS 524
QY 538 TWGAASVTLTVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVWGIKQOARILAVERYLK 597
Db 525 TWGAASVTLTVQARQLLSGIVQOQSNLLRAIEAQHLLQLTVWGIKQOARILAVERYLK 584
QY 598 DQQLLGWCSGKLICTTTVPWNASWSKNTLEAIWNMTMOWDKEIDNTSYLSLIE 657
Db 585 DQQLLGWCSGKLICTTTVPWNASWS-NKSLSDIWNMTWMEWEREIDNTSYLSLIE 643
QY 658 ESPIQOEKNEQELLELDKWANLWNPDISNWLWYIKIFIMIVGVLGRLRVFVLSIVNR 717
Db 644 DSQIQOEKNEQELLELDKWANLWNPDISNWLWYIKIFIMIVGVLGRLRVFVLSIVNR 703
QY 718 VRQGYSPLSFOTRLPAPGDPDRPEIEBEGDRDRSGLLVDGFLTLIWDLSLCLFS 777
Db 704 VRQGYSPLSFOTRLPAPGDPDRPEIEBEGDRDRSGLLVDGFLTLIWDLSLCLFS 763
QY 778 YHRLRDLIIIVTRIVELGRGWEILKYWNLLQVWSOELKNSAVSLFNATAIAVAVGTD 837
Db 764 YHRLRDLIIIVTRIVELGRGWEILKYWNLLQVWSOELKNSAVSLFNATAIAVAVGTD 823
QY 838 RVTIQVQRVGRALLHIPTIRIQLERAL 865
Db 824 RALEVVQRAFRAILHIPTIRIQLERAL 851

RESULT 8
VCLJLV
env polyprotein precursor - human immunodeficiency virus type 1 (isolate LAV-1a)
N:Alternate names: coat polyprotein
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 09-Jul-2004
C:Accession: A03975
R:Wain-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Alizon, M.
Cell 40, 9-17, 1985
A:Title: Nucleotide sequence of the AIDS virus, LAV.
A:Reference number: A90866; MUID:85099333; PMID:2981635
A:Accession: A03975
A:Molecule type: DNA
A:Residues: 1-861 <WAI>
A:Cross-references: UNIPROT:P03377; GB:K02013; NID:g326417; PIDN:AAB59751.1; PID:g326424
C:Genetics:
C:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprotein
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-516/Product: exterior membrane glycoprotein #status predicted <EXT>
F:517-861/Product: transmembrane glycoprotein #status predicted <TM>
F:88,136,141,146,161,165,191,202,239,246,267,281,294,306,337,344,361,391,397,402,411,
F:616,621,630,642,679,755,821/Binding site: carbohydrate (Asn) (covalent) #status predict
Query Match 81.3%; Score 3753.5; DB 1; Length 861;
Best Local Similarity 82.6%; Pred. No. 4.9e-261;
Matches 721; Conservative 51; Mismatches 82; Indels 19; Gaps 11;
```


Db 764 CLFSYHRLDLLLIVTRIVELGRRGWALKYWNLLQYWSQELKNSAVSLNATAVA 823
Qy 834 EGTDRVIVQLVRGVRGALLHIPTIRIQLERALL 866
Db 824 EGTDRVIEVVGAYRAIRHPRIRIQLERILL 856
RESULT 10
VCLJVL
env polyprotein precursor - human immunodeficiency virus type 1 (isolate LV)
N;Alternate names: coat polyprotein
C;Species: human immunodeficiency virus type 1, HIV-1
A;Note: host Homo sapiens (man)
C;Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 09-Jul-2004
C;Accession: A03974
R;Muesing, M.A.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Lasky, L.A.; Capon, D.J.
Nature 313, 450-458, 1985
A;Title: Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus
A;Reference number: A93355; MUID:8511157; PMID:2982104
A;Accession: A03974
A;Molecule type: DNA
A;Residues: 1-856 <MUE>
A;Cross-references: UNIPROT:P03376; GB:K02083; MID:g555008; PIDN:AAB59873.1; PID:g328559
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprotein
F;1-30/Domain: signal sequence #status predicted <SIG>
F;31-511/Product: exterior membrane glycoprotein #status predicted <EXT>
F;512-856/Product: transmembrane glycoprotein #status predicted <TM>
F;88,136,141,156,160,186,197,230,234,241,262,276,289,295,301,332,339,356,386,392,397,406
F;611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 80.8%; Score 3731; DB 1; Length 856;
Best Local Similarity 82.1%; Pred. No. 2e-259;
Matches 717; Conservative 50; Mismatches 82; Indels 24; Gaps 11;
Qy 1 MRVKGIRRYQH---W-WGWTMLLGLMLICSAATEKLWTVVYGVVWKEATTLFCASD 56
Db 1 MRVK---EKYQHLWRGWGWGWTMLLGLMLICSAATEKLWTVVYGVVWKEATTLFCASD 57
Qy 57 AKAYDTEAHNVWATHACVPTDPNPQEVVLNVVTFENFMNKNVWQVQHEHDIISLWQSLK 116
Db 58 AKAYDTEAHNVWATHACVPTDPNPQEVVLNVVTFENFMNKNVWQVQHEHDIISLWQSLK 117
Qy 117 PCVKLTPCLVTLNCTDLRNTNTNNSDNNNSBGTIKGEMKNCNFNATSIGDKMQK 176
Db 118 PCVKLTPCLVSLKCTDLKNDTNT-----NSSSGRMIMKEGEEKNCNFNATSIGDKMQK 171
Qy 177 EYALLYKLDIEPIDNDNTSVRLISCNSTSVITQACPKISPEPIPIHYCAPAGFAILLKCNCK 236
Db 172 EYAFFYKLDIPIIDNDTTSVTLSCNSTSVITQACPKVSPPEPIPIHYCAPAGFAILLKCNCK 231
Qy 237 KPSGKSGCKNVSTVQCTHGIRPVVSTQLLINGSLAEEVIRSENFNFNNAKTIIVQLREP 296
Db 232 TFGTGCTGVSTVQCTHGIRPVVSTQLLINGSLAEEVIRSENFNFNNAKTIIVQLNQS 291
Qy 297 VKINCRPNNTKSI PM--GPGRAFVTTGQIGDIROAHCHNISKNTWTWALKQVVEKLG 354
Db 292 VEINCTRPNNTRKSI RIQKPGRAFVTTGK--IGNMQAHCHNISRANKNTLQKIDSKLR 350
Qy 355 EQFNKTK-IVFTNSSGGDPEIVTHSFNCAGEFFYCNLTQLFDSIWNSENGTWNITRGLNN 413
Db 351 EQFGNKTIIIFKQSSGGDPEIVTHSFNCAGEFFYCNLTQLFNSTW--FNSTWS--TEGSNN 407
Qy 414 TGRNDITLPCRKOIINRQEVGKAMYPPIKGNITSCSNITGLLITRDGGKDDNSRDG 473
Db 408 TEGSDITLPCRKOIINRQEVGKAMYPPIISQICRCSNITGLLITRDGGNNNE---464
Qy 474 NETFRGGGDMRNRSELKYKVKIEPLGVAPTAKRRVVOEERAVGLGAMFFGLG 533
Db 465 SEIFRGGGDMRNRSELKYKVKIEPLGVAPTAKRRVVOEERAVGLGALFLGFLG 524

Qy 534 AAGSTWGAASVTITVQARQLLSGIVQOQSNLLRAIEAQHLLQLTWGIKOLQARILAVE 593
Db 525 AAGSTWGAASMTITVQARQLLSGIVQOQSNLLRAIEAQHLLQLTWGIKOLQARILAVE 584
Qy 594 RYLKDOQLLIGWCSGKLICTTTVPWNASWKNKTELEAIWNNMTWQMDKEIDNNTSLIY 653
Db 585 RYLKDOQLLIGWCSGKLICTTAVPNASWS--NKSLEQIWNNTWMEWDREINNTSLIH 643
Qy 654 SLIEESPIQOEKNEQELLELDKWNANWNPDISNWLWYIKIFIMIVGGLVGLRIVFVVL 713
Db 644 SLIEESNQOEKNEQELLELDKWNANWNPDISNWLWYIKIFIMIVGGLVGLRIVFVVL 703
Qy 714 INVRVQGYSPISFQTRLPAPRGPDRPEETEEGGDRDRDRSGLLVDGFTLIWVDRSL 773
Db 704 INVRVQGYSPISFQTRLPAPRGPDRPEETEEGGDRDRDRSGLLVDGFTLIWVDRSL 763
Qy 774 CLFSYHRLDLLLIVTRIVELGRRGWALKYWNLLQYWSQELKNSAVSLNATAVA 833
Db 764 CLFSYHRLDLLLIVTRIVELGRRGWALKYWNLLQYWSQELKNSAVSLNATAVA 823
Qy 834 EGTDRVIVQLVRGVRGALLHIPTIRIQLERALL 866
Db 824 EGTDRVIEVVGAYRAIRHPRIRIQLERILL 856
RESULT 11
S13288
env protein - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S13288
R;O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A.
Nature 348, 69-73, 1990
A;Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120 c
A;Reference number: S13288; MUID:91043044; PMID:2172833
A;Accession: S13288
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-854 <OBR>
A;Cross-references: UNIPROT:Q85582; UNIPROT:Q72502; UNIPROT:Q09178; UNIPROT:Q78243; UNIP
C;Superfamily: type E retrovirus env polyprotein
Query Match 80.5%; Score 3716; DB 2; Length 854;
Best Local Similarity 81.9%; Pred. No. 2.4e-258;
Matches 715; Conservative 51; Mismatches 28; Indels 26; Gaps 12;
Qy 1 MRVKGIRRYQH---W-WGWTMLLGLMLICSAATEKLWTVVYGVVWKEATTLFCASD 56
Db 1 MRVK---EKYQHLWRGWGWGWTMLLGLMLICSAATEKLWTVVYGVVWKEATTLFCASD 57
Qy 57 AKAYDTEAHNVWATHACVPTDPNPQEVVLNVVTFENFMNKNVWQVQHEHDIISLWQSLK 116
Db 58 AKAYDTEAHNVWATHACVPTDPNPQEVVLNVVTFENFMNKNVWQVQHEHDIISLWQSLK 117
Qy 117 PCVKLTPCLVTLNCTDLRNTNTNNSDNNNSBGTIKGEMKNCNFNATSIGDKMQK 176
Db 118 PCVKLTPCLVSLKCTDLKNDTNT-----NSSSGRMIMKEGEEKNCNFNATSIGDKMQK 171
Qy 177 EYALLYKLDIEPIDNDNTSVRLISCNSTSVITQACPKISPEPIPIHYCAPAGFAILLKCNCK 236
Db 172 EYAFFYKLDIPII--DNTSVRLISCNSTSVITQACPKVSPPEPIPIHYCAPAGFAILLKCNCK 229
Qy 237 KPSGKSGCKNVSTVQCTHGIRPVVSTQLLINGSLAEEVIRSENFNFNNAKTIIVQLREP 296
Db 230 TFGTGCTGVSTVQCTHGIRPVVSTQLLINGSLAEEVIRSENFNFNNAKTIIVQLNQS 289
Qy 297 VKINCRPNNTKSI PM--GPGRAFVTTGQIGDIROAHCHNISKNTWTWALKQVVEKLG 354
Db 290 VEINCTRPNNTRKSI RIQKPGRAFVTTGK--IGNMQAHCHNISRANKNTLQKIDSKLR 348
Qy 355 EQFNKTK-IVFTNSSGGDPEIVTHSFNCAGEFFYCNLTQLFDSIWNSENGTWNITRGLNN 413

Db 349 EQGNKNTIIFKQSSGGDPVIVTHSFNCGGEFFVCNSTQLFNSWTW--FNSWTWS--TEGSNN 405
QY 414 TGRNDTITLPCRKIQKINRWQEVGKAMYAPPIKGNISCSNITGLLTLTRDGGKDDNSRDG 473
Db 406 TEGSDTITLPCRKIQKINRWQEVGKAMYAPPIKGNISCSNITGLLTLTRDGG--NNNG 462
QY 474 NETPRPGGDMRDNRSELKYKVKVIEPLGVAPTAKRRVVOREERAVGLGAMFFGFLG 533
Db 463 SEIFRPGGDMRDNRSELKYKVKVIEPLGVAPTAKRRVVOREERAVGLGAMFFGFLG 522
QY 534 AAGSTWGAASVTLTVQARQLLSGIVQOQNLLRAIEAQHLLQLTWGIKQLOARILAVE 593
Db 523 AAGSTWGAASVTLTVQARQLLSGIVQOQNLLRAIEAQHLLQLTWGIKQLOARILAVE 582
QY 594 RYLKDOQLLGIWCSGKLICTTTVPNWSKNTLEAIWNNMTWOWDKEIDNYSLIY 653
Db 583 RYLKDOQLLGIWCSGKLICTTTVPNWSKNTLEAIWNNMTWOWDKEIDNYSLIH 641
QY 654 SLIEESPIQOEKNEQELLELDKWANLWNNFDSINWLIWYIKIFIMIVGGLVGLRIVFVLS 713
Db 642 SLIEESQOQEKNEQELLELDKWANLWNNFDSINWLIWYIKIFIMIVGGLVGLRIVFVLS 701
QY 714 IVNRVQGYSPLSFQTHLPAPRPGDRPEEIEEGGDRDRSGLLVDDGFTLTLWDLRL 773
Db 702 IVNRVQGYSPLSFQTHLPAPRPGDRPEEIEEGGDRDRSGLLVDDGFTLTLWDLRL 761
QY 774 CLFSYHRLDLLLITVTRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATAI 833
Db 762 CLFSYHRLDLLLITVTRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATAI 821
QY 834 EGTDRVIVQVORVGRALLHPTPIRQGLERALL 866
Db 822 EGTDRVIVQVORVGRALLHPTPIRQGLERALL 854

RESULT 12
S33985
env polyprotein - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S33985
R;Carlini, F.
submitted to the EMBL Data Library, November 1991
A;Reference number: S33979
A;Accession: S33985
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-851 <CAR>
A;Cross-references: UNIPROT:Q78243; EMBL:Z11530; NID:g60192; PIDN:CAA77628.1; PID:g60199
C;Superfamily: type E retrovirus env polyprotein

Query Match 80.4%; Score 3710.5; DB 2; Length 851;
Best Local Similarity 81.7%; Pred. No. 5.9e-258;
Matches 713; Conservative 51; Mismatches 80; Indels 29; Gaps 10;

QY 1 MRVKGIRNTQH---W-WGWTMLLGLLMLTCSATEKLWTVVYGVVPWKAEATTLFCASD 56
Db 1 MRVK---EKYQHLWRWGRWGTMGLLMLTCSATEKLWTVVYGVVPWKAEATTLFCASD 57
QY 57 AKAYDEAHNVWATHACVPTDPNPOEVLNVNTEFNWKNVQWQHEHDIISLDQSLK 116
Db 58 AKAYDEVHNVWATHACVPTDPNPOEVLNVNTEFNWKNVQWQHEHDIISLDQSLK 117
QY 117 PCVKLTPLCVLTLNCTDLRNTNTNNSDNNSEGTIKGEMKNCNFNIATSIGDKWQK 176
Db 118 PCVKLTPLCVLTLNCTDLRNTNTNNSDNNSEGTIKGEMKNCNFNIATSIGDKWQK 171
QY 177 EYALLYKLDIEPDNDNTSTRLSCNTSVITQACPKISFEPIPIHYCAPAGFAILKCNK 236
Db 172 EYAFFYKLDIIPDNDNTSTRLSCNTSVITQACPKISFEPIPIHYCAPAGFAILKCNK 231
QY 237 KFSKSGCKNVSTVQCTHGIRPVVSTQLLNGSLABEEVIRSVNFNNAKTIIVOLREP 296
Db 238 KFSKSGCKNVSTVQCTHGIRPVVSTQLLNGSLABEEVIRSVNFNNAKTIIVOLREP 296

Db 232 TFNGTGPCTNVSTVQCTHGIRPVVSTQLLNGSLABEEVIRSVNFNNAKTIIVOLNTS 291
QY 297 VKINCSPNNNTKRSIPM--GPGRAFVTTCQIICDIRQAHNCISKTNTNATLQVVEKLG 354
Db 292 VEINCTFPNNNTKRIQKIRGPGRAFVTIGK-IGNRQAHNCISRAKWNNTLQKIDSKLR 350
QY 355 EQFNKTK-IVFTNSSGGDPBEIVTHSPNCAGFFPYCNTTQLFDSIWNSENCTWNTGRLNN 413
Db 351 EQFGNNTIIFKQSSGGDPBEIVTHSPNCAGFFPYCNTTQLFDSIWNSENCTWNTGRLNN 402
QY 414 TGRNDTITLPCRKIQKINRWQEVGKAMYAPPIKGNISCSNITGLLTLTRDGGKDDNSRDG 473
Db 403 TEGSDTITLPCRKIQKINRWQEVGKAMYAPPIKGNISCSNITGLLTLTRDGG--NSNNE 459
QY 474 NETPRPGGDMRDNRSELKYKVKVIEPLGVAPTAKRRVVOREERAVGLGAMFFGFLG 533
Db 460 SEIFRPGGDMRDNRSELKYKVKVIEPLGVAPTAKRRVVOREERAVGLGAMFFGFLG 519
QY 534 AAGSTWGAASVTLTVQARQLLSGIVQOQNLLRAIEAQHLLQLTWGIKQLOARILAVE 593
Db 520 AAGSTWGAASVTLTVQARQLLSGIVQOQNLLRAIEAQHLLQLTWGIKQLOARILAVE 579
QY 594 RYLKDOQLLGIWCSGKLICTTTVPNWSKNTLEAIWNNMTWOWDKEIDNYSLIY 653
Db 580 RYLKDOQLLGIWCSGKLICTTTVPNWSKNTLEAIWNNMTWOWDKEIDNYSLIH 638
QY 654 SLIEESPIQOEKNEQELLELDKWANLWNNFDSINWLIWYIKIFIMIVGGLVGLRIVFVLS 713
Db 639 SLIEESQOQEKNEQELLELDKWANLWNNFDSINWLIWYIKIFIMIVGGLVGLRIVFVLS 698
QY 714 IVNRVQGYSPLSFQTHLPAPRPGDRPEEIEEGGDRDRSGLLVDDGFTLTLWDLRL 773
Db 699 IVNRVQGYSPLSFQTHLPAPRPGDRPEEIEEGGDRDRSGLLVDDGFTLTLWDLRL 758
QY 774 CLFSYHRLDLLLITVTRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATAI 833
Db 759 CLFSYHRLDLLLITVTRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATAI 818
QY 834 EGTDRVIVQVORVGRALLHPTPIRQGLERALL 866
Db 819 EGTDRVIVQVORVGRALLHPTPIRQGLERALL 851

RESULT 13
VCLJH4
env polyprotein - human immunodeficiency virus type 1 (isolate CDC-451)
N;Alternate names: coat polyprotein
N;Contains: coat protein gp120; coat protein gp41
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 09-Jul-2004
C;Accession: C25523
R;Desai, S.M.; Kalyanaraman, V.S.; Casey, J.M.; Srinivasan, A.; Andersen, P.R.; Devare, S.
Proc. Natl. Acad. Sci. U.S.A. 83, 8380-8384, 1986
A;Title: Molecular cloning and primary nucleotide sequence analysis of a distinct human
A;Reference number: A94136; MUID:87041461; PMID:3490666
A;Accession: C25523
A;Molecule type: DNA
A;Residues: 1-868 <DES>
A;Cross-references: UNIPROT:P05879; GB:M13137; NID:g326460; PIDN:AAA44311.1; PID:g326467
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane protei
F;1-521/Product: coat protein gp120 #status predicted <GP1>
F;522-868/Product: coat protein gp41 #status predicted <GP2>
F;89,138,139,142,166,195,198,208,245,252,273,287,300,312,342,349,365,371,395,405,409,459,
Query Match 80.3%; Score 3708; DB 1; Length 868;
Best Local Similarity 80.5%; Pred. No. 9.2e-258;
Matches 703; Conservative 56; Mismatches 100; Indels 14; Gaps 7;

QY 1 MRVKGIRNTQH---W-WGWTMLLGLLMLTCSATEKLWTVVYGVVPWKAEATTLFCASD 60
Db 1 MRVKGIRNTQH---W-WGWTMLLGLLMLTCSATEKLWTVVYGVVPWKAEATTLFCASD 60

Db 3 MRAKIGIRKNCQHLRWGTMLLGMLMCSAAANLWVTVYVYGVVWKEATTTLFCASDAKAY 62
QY 61 DTAHNWATHACVPTDNPQVELVNVNFTNFMWKNMVEQMHEDIISLWDQSLKPCVK 120
Db 63 DTAHNWATHACVPTNPNQVELNVNFTNFMWKNMVEQMHEDIISLWDQSLKPCVK 122
QY 121 LTPCLVTLNCTDLRNTTNTNNSSEGTIKGEMKNCNFNATSTGDKWQKEYAL 180
Db 123 LTPCLVTLNCTDLRNTTNTTNTLSLIVVWQGGKGMKNCNFNATSTIRDKQREYAL 181
QY 181 LYKLDIEPIDN-----DNTSYRLISCNSTVITQACPKISPEPIPIHYCAPAGFAILKND 235
Db 182 FYKLDVEPIDDKNTTNTKYRLJNCNTSVITQACPKVSEPIPIHYCTPTGFALLKND 241
QY 236 KPSGKSGCKNVSTQCTHGIRPVSTQLLNGSLAEVEVIRSENFNTNAKTIIVOLRE 295
Db 242 KFNKGTPCTNVSTQCTHGIRPVSTQLLNGSLAEVEVIRSENFNTNAKTIIVQLNV 301
QY 296 PVKINCSRPNNTRKSPMCPGAFYTTGQIIGDIRQAHCHNISKTWNTWALKOVVEKGE 355
Db 302 SVBINCTRPNHTRKRVTLGPGRWVYTTGELGNIRQAHCHNISRAQWNTLQOIATTLRE 361
QY 356 QFNKTKIVFTNSGGDPEIVTHSFNCAGEFFYCNVTTQLFDSIWN-SENGTWNITRGLNNT 414
Db 362 QFNKTKIAFNQSGDPEIVWHSFNCGEFFYCNVTFNSANVTNSGTWSTVRKQDT 421
QY 415 GRNDTITLPCRIKOIINRWQEVGKAMYAPPIKGNISCNSTGLLITRDGKDDNSRDGN 474
Db 422 G--DIITLPCRIKOIINRWQGVGKAMYPKGLIRCSSNITGLLITRDGCGENQT---T 476
QY 475 ETRPFGGDMRDNRSELYKYVVKIEPLGVAPTAKARRVQREERAVG-LGAMFFGFLG 533
Db 477 ETRPFGGDMRDNRSELYKYVVKIEPLGVAPTAKARRVQREERAVGMLGAMFFGFLG 536
QY 534 AAGSTWGAASVTLTVQARQLLSGIVQOOSNLLRAIEAQHLLQLTVMGIKQOLQARILAVE 593
Db 537 AAGSTWGAASVTLTVQARQLLSGIVQOOSNLLRAIEAQHLLQLTVMGIKQOLQARILAVE 596
QY 594 RYLDQQLLGGVCSGKLCITTTVPWNASWKNKTLLEAIWNNMTWQMDKEIDNYTSLIY 653
Db 597 RYLDQQLLGGVCSGKLCITTTVPWNASWKNKTLLEAIWNNMTWQMDKEIDNYTSLIY 655
QY 654 SLIEEPIQEKNEQELLEDKLANLWNVFDISNWLWIKIPIMIVGGLVGLIRIVFVLS 713
Db 656 TLIEESQNEQEKNEQELLEDKLANLWNVFDISNWLWIKIPIMIVGGLVGLIRIVFVLS 715
QY 714 INVRVQGYSPSLFOTLPAPRGDPEETEEBGGDRDRSGLLVDGFLTLWVDLRLS 773
Db 716 INVRVQGYSPSLFOTLPAPRGDPEETEEBGGDRDRSGLLVDGFLTLWVDDRLS 775
QY 774 CLFSYHRLDLLITVTRIVELLGRRGWELIKYWNLLQYWSQELKNSAVSLFNATAIYA 833
Db 776 CLFSYHRLDLLITVTRIVELLGRRGWELIKYWNLLQYWSQELKNSAVSLFNATAIYA 835
QY 834 EGTDRVIVLQVRGALLHPTRIROGLERALL 866
Db 836 EGTDRVIVLQVRGALLHPTRIROGLERALL 868
RESULT 14
vcljkb
env polyprotein precursor - human immunodeficiency virus type 1 (strain KB-1-gp41)
N;Alternate names: coat polyprotein
N;Contigins: coat protein gp120; coat protein gp41
C;Species: human immunodeficiency virus type 1, HIV-1
A;Note: host Homo sapiens (man)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 09-Jul-2004
C;Accession: A42995
R;Shimizu, H.; Hasebe, F.; Tsuchie, H.; Morikawa, S.; Ushijima, H.; Kitamura, T.
Virology 189, 534-546, 1992
A;Title: Analysis of a human immunodeficiency virus type 1 isolate carrying a truncated
A;Reference number: A42995; MUID:92351552; PMID:1322587
A;Accession: A42995

A;Molecule type: mRNA
A;Residues: 1-861 <SH1>
A;Cross-references: UNIPROT:P31819; GB:S41266; GB:D01206
C;Genetics:
A;Gene: env
C;Superfamily: type B retrovirus env polyprotein
C;Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F;1-689/Domain: extracellular #status predicted <EXT>
F;17-33/Domain: signal sequence #status predicted <SIG>
F;17-33/Region: hydrophobic #status predicted
F;34-517/Product: coat protein gp120 #status predicted <CP1>
F;514-517/Region: cleavage processing #status predicted
F;518-861/Product: coat protein gp41 #status predicted <CP2>
F;518-534/Region: hydrophobic #status predicted
F;590-711/Domain: transmembrane #status predicted <TM1>
F;712-861/Domain: intracellular #status predicted <INT>
F;756-772/Region: hydrophobic #status predicted
F;93,141,145,146,163,191,192,237,241,248,269,283,296,308,338,345,361,367,397,403,408,414,
Query Match 80.2%; Score 3703.5; DB 1; Length 861;
Best Local Similarity 80.2%; Pred. No. 1.9e-257;
Matches 704; Conservative 65; Mismatches 80; Indels 29; Gaps 12;
QY 1 MRVKGIRRNQOH-W-WG-----WGTMLLGLMICSATEKLMVTVYVYGVVWKEATTTLFCFA 54
Db 1 MRVKEIRKYOHLRWGIMLRWGTMLLGLMICSAAEQLWVTVYVYGVVWKEATTTLFCFA 60
QY 55 SDAKAYDTEAHNWNATHACVPTDNPQVELNVNFTNFMWKNMVEQMHEDIISLWDQS 114
Db 61 SDAKAYDTEAHNWNATHACVPTDNPQVELNVNFTNFMWKNMVEQMHEDIISLWDQS 120
QY 115 LKPCVKLPLCVTLNCTDLRNTTNTNNSSEGTIKGEMKNCNFNATSTIGDKM 174
Db 121 LKPCVKLPLCVTLNCTDLRNTT-----NNSSIEEKMK-GEIKNCSNFVNTNRDKV 172
QY 175 QKEYALLKLDIEPIDN-----NTSVRLISCNSTSVITQACPKISPEPIHYCAPAGFAI 230
Db 173 QKEYALLKLDIEPIDN-----NTSVRLISCNSTSVITQACPKISPEPIHYCAPAGFAI 232
QY 231 LKCNCKFKSGKCKNVSTQCTHGIRPVSTQLLNGSLAEVEVIRSENFNTNAKTI 290
Db 233 LKCNCKFKSGKCKNVSTQCTHGIRPVSTQLLNGSLAEVEVIRSENFNTNAKTI 292
QY 291 VQLREPVKINCSRPNNTRKSPMCPGAFYTTGQIIGDIRQAHCHNISKTWNTWALKOVV 350
Db 293 VQLNETVKNICIRPNKTRKRVTMGPGRVYTTGEEIIGDIRQAHCHNISAEWNTLEQIA 352
QY 351 EKLGEQFNKTKIVFTNSGGDPEIVTHSFNCAGEFFYCNVTTQLFDSIWNSENGTWNITRG 410
Db 353 NKLRQFENKTIIVFNOSGGDPEIVMHNFCGGEFFYCDSSQLFNST-HLSNGTW----- 406
QY 411 LNNTGRNDTITLPCRIKOIINRWQEVGKAMYAPPIKGNISCNSTGLLITRDGG--KDD 468
Db 407 WNGTG-PENITLPCRIKOIINRWQEVGKAMYAPPIRQIRCSSNITGLLITRDGGNTQNN 465
QY 469 NSRDGNETRPPGGDMRDNRSELYKYVVKIEPLGVAPTAKARRVQREERAVGLGAMF 528
Db 466 NTSNSEEIRPPGGDMRDNRSELYKYVVKIEPLGVAPTAKARRVQREERAVGLGAMF 525
QY 529 FGLGAGSTWGAASVTLTVQARQLLSGIVQOOSNLLRAIEAQHLLQLTVMGIKQOLQAR 588
Db 526 LGLGAGSTWGAASVTLTVQARQLLSGIVQOOSNLLRAIEAQHLLQLTVMGIKQOLQAR 585
QY 589 ILAVERYLQDQQLLGIWCGSKLCITTTVPWNASWKNKTLLEAIWNNMTWQMDKEIDNY 648
Db 586 VLAVERYLQDQQLLGIWCGSKLCITTTVPWNTWS-NKSFNEIWNMTWMEWEIREINNY 644
QY 649 TSLIYSLIEEPIQEKNEQELLEDKLANLWNVFDISNWLWIKIPIMIVGGLVGLIRIV 708
Db 645 TNLINLIEEESQNEQEKNEQELLEDKLANLWNVFDISNWLWIKIPIMIVGGLVGLIRIV 704
QY 709 FVVLISVNRVQGYSPSLFOTLPAPRGDPEETEEBGGDRDRSGLLVDGFLTLWV 768
Db 709 FVVLISVNRVQGYSPSLFOTLPAPRGDPEETEEBGGDRDRSGLLVDGFLTLWV 768

Db 705 FTVLSIVNRVQGYSPUSFQTRLPA-RGDRPRGIEEGGERDRDRSGPLVDGLLAIWV 763
QY 769 DLRLCLFSYHRLRDLILLIVTRIVELLGRGWELTKYWNLLQYWSQELKNSAVSLFNAT 828
Db 764 DLRLCLFSYHRLRDLILLIVTRIVELLGRGWELTKYWNLLQYWSQELKNSAVSLFNAT 823
QY 829 AIAVAEGTDRVIOVLQVRGRLHLLHIPRIRQGLERALL 866
Db 824 AIAVAEGTDRVIEQLQYRAILLHIPVKIRQGLERALL 861

RESULT 15
VCLJ3W
env polyprotein precursor - human immunodeficiency virus type 1 (isolate WMJ1)
N:Contains: coat protein gp120; coat protein gp41
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 09-Jul-2004
C:Accession: A24774
R:Starcich, B.R.; Hahn, B.H.; Shaw, G.M.; McNeely, P.D.; Modrow, S.; Wolf, H.; Parks, E.
Cell 45, 637-648, 1986
A:Title: Identification and characterization of conserved and variable regions in the en
A:Reference number: A24774; MUID:86218077; PMID:2423250
A:Accession: A24774
A:Molecule type: DNA
A:Residues: 1-856 <STA>
A:Cross-references: UNIPROT:P31872; GB:K03455; GB:M38432; NID:gi906382
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-501/Product: coat protein gp120 #status predicted <GP1>
F:502-847/Product: coat protein gp41 #status predicted <GP2>
F:87,134,140,151,155,183,197,234,241,262,276,289,295,331,338,354,360,390,394,404,447,459

Query Match 80.1%; Score 3700; DB 1; Length 856;
Best Local Similarity 81.5%; Pred. No. 3.4e-257;
Matches 712; Conservative 49; Mismatches 87; Indels 26; Gaps 10;

QY 1 MRVKGIRRNQHWGWTMLLGLLMICSATEKLMWTVYGVVPMKCATTTLFCASDAKAY 60
Db 1 MRVKGIRRNQHLWIWGTMLFGWMMICSATVEQLWTVYGVVPMKCATTTLFCASDAKAY 60
QY 61 DTEAHNVWATHACVPDTPNPQVELNVVTENFNWKNMVEQMHEDIISLWDQSLKPCVK 120
Db 61 STEAHKWATHACVPNPFPQEVLENVTENFNWKNMVEQMHEDIISLWDQSLKPCVK 120
QY 121 LTPLCVTNLCTDLRNTNTNNSDNNNSSEGIKGMKNCSPNATSGDKMOKEYAL 180
Db 121 LTPLCVTNLCTDLRNTNTNNSDNNNSSEGIKGMKNCSPNATSGDKMOKEYAL 170
QY 181 LYKLDIEPIDNDN----TSYRLISNTSVITQACPKIPHYCAPAGFAILKND 235
Db 171 FYKLDVVPKSNDSSTYTRYLIHNTSVITQACSKVSFEPIHYCAPAGFAILKND 230
QY 236 KKFSGKSCKNSTVQCTHGIRPVWSTQLLNGSLAEEVIRSENFTNNAKTIIVQLRE 295
Db 231 KKFNGTGPCTNVSTVQCTHGIRPVWSTQLLNGSLAEEVIRSENFTNNAKTIIVHLNE 290
QY 296 PVKINCSRPNNNT-RKSIIPMGPGAFYTTQIIGDTRQAHNCISKYNWTNALKQVVEKLG 354
Db 291 SVEINCTRPNNNVRRRIHIGPGAFY-TGEIRGNTRQAHNCISRAKWNNTLKQIVEKLR 349
QY 355 EQFNKTKIVFTNSGGDPEIVTHSFNCAGFEFFYCNITQLFDSIWNSENGTWNITRGLNNT 414
Db 350 EQFNKTIIVNHSGGDPEIVTHSFNCGGFEFFYCDSTQLFNSTWV---TGISTEGNNT 406
QY 415 GRN-DTITLPCRKIQRNMQVQKAMYPPIKGNISCSNITGLLLTRDGGKDDNSRDG 473
Db 407 EENGDTITLPCRKIQRNMQVQKAMYPPIKGNISCSNITGLLLTRDGG---NSSSR 463
QY 474 NETFRGGDMRNRSSELYKVKVKEPLGVAPTAKRRVQREBRANG-LGAMFPGFL 532

Db 464 EEIFRPGGNMRDNRSSELYKVKVKEPLGVAPTAKRRVQREBRANGAIGAMFLGFL 523
QY 533 GAAGSTMGAASVTLTYQARQLLSGIVQOQSNLIRAJEAQOHLQLTWGIGKQLQARILAV 592
Db 524 GAAGSTMGAASVTLTYQARQLLSGIVQOQSNLIRAJEAQOHLQLTWGIGKQLQARILAV 593
QY 593 ERYLKDQQLLGIWGCSSGKLICTTTVPWNASWSKNKTLEAIWNNMTWQMDKEIDNVTSLI 652
Db 584 ERYLRDQQLLGIWGCSSGKLICTTTVPWNASWS-NKSMQDIWNNMTWMEWEREIDNVTSLI 642
QY 653 YSLIEESPQOEKNEQELLEDKWNINWFDISNMLWYIKIFIMIVGGVGLVIRIVFVVL 712
Db 643 YNLIEESQOEKNEQELLEDKWNINWFDISNMLWYIKIFIMIVGGVGLVIRIVFVSVL 702
QY 713 SIVNRVQGYSPUSFQTRLPA-RGDRPRGIEEGGERDRDRSGPLVDGLLAIWV 772
Db 703 SIVNRVQGYSPUSFQTRLPA-RGDRPRGIEEGGERDRDRSGPLVDGLLAIWV 762
QY 773 LCLFSYHRLRDLILLIVTRIVELLGRGWELTKYWNLLQYWSQELKNSAVSLFNAT 832
Db 763 LCLFSYHRLRDLILLIVTRIVELLGRGWELTKYWNLLQYWSQELKNSAVSLFNAT 822
QY 833 AEGTDRVIOVLQVRGRLHLLHIPRIRQGLERALL 866
Db 823 AEGTDRVIEVQVQKICRAIIHIPRIRQGLERALL 856

Search completed: April 11, 2005, 11:15:28
Job time : 77.6201 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2005, 10:39:11 ; Search time 277.268 Seconds
(without alignments)
1599.394 Million cell updates/sec

Title: US-09-762-261C-1

Perfect score: 4617

Sequence: 1 MRVKGIRRYQHWWGWTML.....GRALLHIPRIQGLERALL 866

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4600	99.6	866	2 Q9WP24	Q9WP24 human immun
2	3990.5	86.4	861	2 Q90294	Q90294 human immun
3	3986.5	86.3	857	2 Q6BC02	Q6BC02 human immun
4	3970	86.0	856	1 ENV_HVIMN	P05877 human immun
5	3969.5	86.0	857	2 Q9YU15	Q9YU15 human immun
6	3964.5	85.9	861	2 Q90295	Q90295 human immun
7	3941.5	85.4	857	2 Q90296	Q90296 human immun
8	3931.5	85.2	853	2 Q7SVL6	Q7SVL6 human immun
9	3928	85.1	850	2 Q7SVL4	Q7SVL4 human immun
10	3917.5	84.8	855	2 Q7SVL0	Q7SVL0 human immun
11	3905	84.6	860	2 Q7SVL7	Q7SVL7 human immun
12	3896.5	84.4	863	2 Q92763	Q92763 human immun
13	3886	84.2	852	2 Q7SVL5	Q7SVL5 human immun
14	3886	84.2	854	2 Q40222	Q40222 human immun
15	3870.5	83.8	853	2 Q03811	Q03811 human immun
16	3868	83.8	865	2 Q8Q2X5	Q8Q2X5 human immun
17	3863.5	83.7	863	2 Q91U20	Q91U20 human immun
18	3862.5	83.7	863	2 Q91JY9	Q91JY9 human immun
19	3862	83.6	852	2 Q41883	Q41883 human immun
20	3859.5	83.6	863	2 Q8UNX2	Q8UNX2 human immun
21	3859	83.6	852	2 Q92761	Q92761 human immun
22	3857.5	83.5	863	2 Q91JY8	Q91JY8 human immun
23	3856.5	83.5	859	2 Q71260	Q71260 human immun
24	3854	83.5	866	2 Q91K01	Q91K01 human immun
25	3851.5	83.4	863	2 Q91JY7	Q91JY7 human immun
26	3848.5	83.4	859	2 Q7SVL2	Q7SVL2 human immun
27	3848	83.3	866	2 Q6EG58	Q6EG58 human immun
28	3848	83.3	866	2 Q6EK29	Q6EK29 human immun
29	3847.5	83.3	871	2 Q91K13	Q91K13 human immun
30	3846	83.3	858	2 Q6U239	Q6U239 human immun
31	3845.5	83.3	847	2 Q6BC19	Q6BC19 human immun

32	3845.5	83.3	847	2 Q75760	Q75760 human immun
33	3843	83.2	866	2 Q6EJY9	Q6EJY9 human immun
34	3842	83.2	850	2 Q6UYQ1	Q6UYQ1 human immun
35	3840	83.2	870	2 Q8Q2X1	Q8Q2X1 human immun
36	3839.5	83.2	843	1 ENV_HV1Y2	P35961 human immun
37	3838.5	83.1	853	2 Q6UYR2	Q6UYR2 human immun
38	3838.5	83.1	871	2 Q91K14	Q91K14 human immun
39	3838	83.1	866	2 Q6EK45	Q6EK45 human immun
40	3834	83.0	866	2 Q6EK37	Q6EK37 human immun
41	3833	83.0	852	2 Q6UZ34	Q6UZ34 human immun
42	3833	83.0	866	2 Q6EJZ7	Q6EJZ7 human immun
43	3832.5	83.0	849	2 Q6UYQ5	Q6UYQ5 human immun
44	3832	83.0	852	2 Q41885	Q41885 human immun
45	3832	83.0	866	2 Q6EG67	Q6EG67 human immun

ALIGNMENTS

RESULT 1
Q9WP24 ID Q9WP24 PRELIMINARY; PRT; 866 AA.
AC Q9WP24;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Envelope protein.
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP MEDLINE=992261533; PubMed=10221533; DOI=10.1089/08922299311088;
RA Quinlan G.V. Jr., Zhang P.F., Fu D.W., Dong M., Alter H.J.;
RT "Expression and characterization of HIV type 1 envelope protein associated with a broadly reactive neutralizing antibody response.";
RN AIDS Res. Hum. Retroviruses 15:561-570(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Quinlan G.V. Jr., Zhang P.F., Fu D.W., Dong M., Alter H.J.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF128126; AAD40637.3; -;
DR HSSP; P04578; 1DLB.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 866 AA; 98113 MW; 6E76021833F2EACD CRC64;

Query Match 99.6%; Score 4600; DB 2; Length 866;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 863; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MRVKGIRRYQHWWGWTMLGLLMI CSATEKLWVTYVYGVVWKATTTLFCASDAKAY 60

|||||

1 MRVKGIRRYQHWWGWTMLGLLMI CSATEKLWVTYVYGVVWKATTTLFCASDAKAY 60

|||||

61 DTEAHNVWATHACVPTDPNPQVELNVNTEFNWKNMVEQMHEDIISLWDSLKPCVK 120

|||||

61 DTEAHNVWATHACVPTDPNPQVELNVNTEFNWKNMVEQMHEDIISLWDSLKPCVK 120

|||||

121 LTPCLCVTLNCTDLRNTNTNTNNNSSEGTIGKGMKNCNFNATSTGDKMQKEYAL 180

|||||

121 LTPCLCVTLNCTDLRNTNTNTNNNSSEGTIGKGMKNCNFNATSTGDKMQKEYAL 180

|||||

181 LYKLDIEPDNDNTSYRLISCNSTSVITQACPKISFEPIPIHYCAPAGFAILKCNCKKFSG 240

Db 181 LYKLDIEPIDNDNTSVRLISCNSTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG 240
Qy 241 KGSCKNVSTVQCTHGRPVVSTVSTOLLNGLSLAEVEVIRSENFTNNAKTIIVQLREPVKIN 300
Db 241 KGSCKNVSTVQCTHGRPVVSTVSTOLLNGLSLAEVEVIRSENFTNNAKTIIVQLREPVKIN 300
Qy 301 CSRPNNNTKRSIPMGPGRAFYTTGQIIGDIRQAHNCISKTNTWLNALQVVEKLGEQFNKT 360
Db 301 CSRPNNNTKRSIPMGPGRAFYTTGQIIGDIRQAHNCISKTNTWLNALQVVEKLGEQFNKT 360
Qy 361 KIVFTNSSGGDPEIVTHSFNCAGEFFYCNVTQQLFDSIWNSENGTWNITRGLNNTGRNDTI 420
Db 361 KIVFTNSSGGDPEIVTHSFNCAGEFFYCNVTQQLFDSIWNSENGTWNITRGLNNTGRNDTI 420
Qy 421 TLPCKRIQIINRWQEVGKAMYPPIKGNISCSNITGLLLTRDGGKDDNSRDGNETFRPG 480
Db 421 TLPCKRIQIINRWQEVGKAMYPPIKGNISCSNITGLLLTRDGGKDDNSRDGNETFRPG 480
Qy 481 GGDMDNRNRSSELYKYKVKVIEPLGVAPTAKRRVVOREERAVGLGAMFGFLGAAGSTMG 540
Db 481 GGDMDNRNRSSELYKYKVKVIEPLGVAPTAKRRVVOREERAVGLGAMFGFLGAAGSTMG 540
Qy 541 AASVTLTVQARQLLSGIVQOQSNLLRAIEAQOHLQLTVWGIKQLOARILAVERYLKDOQ 600
Db 541 AASVTLTVQARQLLSGIVQOQSNLLRAIEAQOHLQLTVWGIKQLOARILAVERYLKDOQ 600
Qy 601 LLGIWGCCKGLICTTTPVWNASKNKTLEAIWNNTWQDKEIDNYSLIYSLIEESP 660
Db 601 LLGIWGCCKGLICTTTPVWNASKNKTLEAIWNNTWQDKEIDNYSLIYSLIEESP 660
Qy 661 IQQEKNEQELLELDKWNALNWNFDISNWLWYIKIFIMIVGGLVGLRIVFVVLISVNRVQ 720
Db 661 IQQEKNEQELLELDKWNALNWNFDISNWLWYIKIFIMIVGGLVGLRIVFVVLISVNRVQ 720
Qy 721 GYSPLSFQTLRPPAPGDRPEEBEGGDRDRSGLLVDGFTLTIWDLRSICLSFYHR 780
Db 721 GYSPLSFQTLRPPAPGDRPEEBEGGDRDRSGLLVDGFTLTIWDLRSICLSFYHR 780
Qy 781 LRDLILLIVTRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATAIAVAEGTDRVI 840
Db 781 LRDLILLIVTRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATAIAVAEGTDRVI 840
Qy 841 QVLQVRGALLHIPTIRIOGLERALL 866
Db 841 QVLQVRGALLHIPTIRIOGLERALL 866

RESULT 2
O90294 ID O90294 PRELIMINARY; PRT; 861 AA.
AC O90294;
DT 01-NOV-1998 (TEMBLrel. 08, Created)
DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)
DE 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98362112; Pubmed=9696803;
RA Park E.J., Vujcic L.K., Anand R., Theodore T.S., Quinnan G.V. Jr.;
RT "Mutations in both gp120 and gp41 are responsible for the broad
RT neutralization resistance of variant human immunodeficiency virus type
RL 1 MN to antibodies directed at V3 and non-V3 epitopes.";
RL J. Virol. 72:7099-7107(1998).
DR EMBL; AF075719; AAC33788.1; .
DR HSP; P05877; INJ0
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;
KW Transmembrane.
SQ SEQUENCE 861 AA; 97670 MW; 8136732E44445E56 CRC64;
Query Match 86.4%; Score 3990.5; DB 2; Length 861;
Best Local Similarity 86.8%; Pred. No. 8.2e-287;
Matches 752; Conservative 38; Mismatches 67; Indels 9; Gaps 5;
Qy 1 MRVKGIRRYQHWWGWTMLLGLLMICSAATEKLWTVVYGVVPWKAEATTLFCASDAKAY 60
Db 1 MRVKGIRRYQHWWGWTMLLGLLMICSAATEKLWTVVYGVVPWKAEATTLFCASDAKAY 60
Qy 61 DTEAHNVWATHACVPTDNPQVELVNVTFENFMKNMKNVVEQMHEDIISLWQSLKPCVK 120
Db 61 DTEAHNVWATHACVPTDNPQVELVNVTFENFMKNMKNVVEQMHEDIISLWQSLKPCVK 120
Qy 121 LTPLCVTLNCTDLRNTNTNNSTDDNNNSSEGTIKGEMKNCNFNIATSIGDKMKQKEYAL 180
Db 121 LTPLCVTLNCTDLRNTNTNNSTDDNNNSSEGTIKGEMKNCNFNIATSIGDKMKQKEYAL 180
Qy 181 LYKLDIEPIDNDNTSVRLISCNSTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG 240
Db 181 LYKLDIEPIDNDNTSVRLISCNSTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG 240
Qy 241 KGSCKNVSTVQCTHGRPVVSTVSTOLLNGLSLAEVEVIRSENFTNNAKTIIVQLREPVKIN 300
Db 241 KGSCKNVSTVQCTHGRPVVSTVSTOLLNGLSLAEVEVIRSENFTNNAKTIIVQLREPVKIN 300
Qy 301 CSRPNNNTKRSIPMGPGRAFYTTGQIIGDIRQAHNCISKTNTWLNALQVVEKLGEQFNKT 360
Db 301 CSRPNNNTKRSIPMGPGRAFYTTGQIIGDIRQAHNCISKTNTWLNALQVVEKLGEQFNKT 360
Qy 361 KIVFTNSSGGDPEIVTHSFNCAGEFFYCNVTQQLFDSIWNSENGTWNITRGLNNTGRNDTI 420
Db 361 KIVFTNSSGGDPEIVTHSFNCAGEFFYCNVTQQLFDSIWNSENGTWNITRGLNNTGRNDTI 420
Qy 421 TLPCKRIQIINRWQEVGKAMYPPIKGNISCSNITGLLLTRDGGKDDNSRDGNETFRPG 480
Db 421 TLPCKRIQIINRWQEVGKAMYPPIKGNISCSNITGLLLTRDGGKDDNSRDGNETFRPG 480
Qy 481 GGDMDNRNRSSELYKYKVKVIEPLGVAPTAKRRVVOREERAVGLGAMFGFLGAAGSTMG 540
Db 481 GGDMDNRNRSSELYKYKVKVIEPLGVAPTAKRRVVOREERAVGLGAMFGFLGAAGSTMG 540
Qy 541 AASVTLTVQARQLLSGIVQOQSNLLRAIEAQOHLQLTVWGIKQLOARILAVERYLKDOQ 600
Db 541 AASVTLTVQARQLLSGIVQOQSNLLRAIEAQOHLQLTVWGIKQLOARILAVERYLKDOQ 600
Qy 601 LLGIWGCCKGLICTTTPVWNASKNKTLEAIWNNTWQDKEIDNYSLIYSLIEESP 660
Db 601 LLGIWGCCKGLICTTTPVWNASKNKTLEAIWNNTWQDKEIDNYSLIYSLIEESP 660
Qy 661 IQQEKNEQELLELDKWNALNWNFDISNWLWYIKIFIMIVGGLVGLRIVFVVLISVNRVQ 720
Db 661 IQQEKNEQELLELDKWNALNWNFDISNWLWYIKIFIMIVGGLVGLRIVFVVLISVNRVQ 720
Qy 721 GYSPLSFQTLRPPAPGDRPEEBEGGDRDRSGLLVDGFTLTIWDLRSICLSFYHR 780
Db 721 GYSPLSFQTLRPPAPGDRPEEBEGGDRDRSGLLVDGFTLTIWDLRSICLSFYHR 780
Qy 781 LRDLILLIVTRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATAIAVAEGTDRVI 840
Db 781 LRDLILLIVTRIVELLGRGWEILKYWNLLQYWSQELKNSAVSLFNATAIAVAEGTDRVI 840
Qy 841 QVLQVRGALLHIPTIRIOGLERALL 866
Db 841 QVLQVRGALLHIPTIRIOGLERALL 866

RESULT 3
 Q6BC02 PRELIMINARY; PRT; 857 AA.
 ID Q6BC02
 AC Q6BC02; 25-OCT-2004 (T-EMBLrel. 28, Created)
 DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
 DE Envelope glycoprotein (Fragment).
 GN Name=env;
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 [1]
 SEQUENCE FROM N.A.
 RA Binley J.M., Wrin T., Zwick M.B., Korber B., Wang M., Chappey C.,
 Stiegler G., Kunert R., Zolla-Pazner S., Katinger H.,
 Petropoulos C.J., Burton D.R.;
 RT "Comprehensive analysis of the cross-clade neutralizing activity of a
 panel of anti-HIV-1 monoclonal antibodies.";
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY669737; AAT67509.1;
 DR GO; GO:0019031; C:Viral envelope; IEA.
 DR GO; GO:0005198; F:Structural molecule activity; IEA.
 DR InterPro; IPR000328; Env_GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;
 Transmembrane.
 FT NON_TER 857 857
 SQ SEQUENCE 857 AA; 97206 MW; 4ED03FD85C91205 CRC64;
 Query Match 86.3%; Score 3986.5; DB 2; Length 857;
 Best Local Similarity 86.7%; Pred. No. 1.6e-286;
 Matches 751; Conservative 38; Mismatches 68; Indels 9; Gaps 5;
 QY 1 MRVKGIRRNQHWGWTMLLGLLMTCATEKLWTVVYGVVPVKATTLFCASDAKAY 60
 DB 1 MRVKGIRRNQHWGWTMLLGLLMTCATEKLWTVVYGVVPVKATTLFCASDAKAY 60
 QY 61 DTAHNVWATHACVPTDPNPQVELNVNTFNWKNMVEQHEHDIISLDQSLKPCVK 120
 DB 61 DTEHNVWATHACVPTDPNPQVELNVNTFNWKNMVEQHEHDIISLDQSLKPCVK 120
 QY 121 LTPCLVTNCTDLRNTNTNNSNDSGTTKGMKNCSEFNATISIGDKMQKEYAL 180
 DB 121 LTPCLVTNCTDLRNTNTNNSNDSGTTKGMKNCSEFNATISIGDKMQKEYAL 180
 QY 181 LYKLDIEPIDNDNTSYRLISNTSVITQACPKISFEPPIHYCAPAGFAILKCNKKFSG 240
 DB 181 LYKLDIVSIDNDSTSYRLISNTSVITQACPKISFEPPIHYCAPAGFAILKCNKKFSG 240
 QY 241 KGSCKNVSTVQCTHGRPVVSTOLLNGSLAEVIRSENFNTNNAKTIIVQLRPVKIN 300
 DB 241 KGSCKNVSTVQCTHGRPVVSTOLLNGSLAEVIRSENFNTNNAKTIIVHLNVSQIN 300
 QY 301 CSRPNNNTRKSIPIMGGRAPFTYTGIIIGDTRQAHCMISKTNWTKLVQVVEKLGEOFNKT 360
 DB 301 CTRPNNNTRKRIHIGGRAPFTYTGIIIGDTRQAHCMINRAKWNDRQIVSKLKEBQFNK 360
 QY 361 KIVFTNSSGDDPEIVTHSFNCAGEFFYCNNTQTFDSIWNSENGTWNITRGLNNTGRNDTI 420
 DB 361 TIVFNQSSGDDPEIVHVSFNCAGEFFYCNNTPLFNSTWNG-NNTWN----NITGSNNNI 414
 QY 421 TLPCKRIQIINRQEVGKAMVAPIKNISSCNITGLLTRDGGKDNDRDGNFTFRPG 480
 DB 415 TLQCKIKQIINRQEVGKAMVAPIEQIRCSNITGLLTRDGGKDTDND-TEIFRPG 473
 QY 481 GGDWRNWRSELYKYVKVTEPLGVAPTAKRRVQREERAVGLGAMFGFLGAAGSTMG 540
 DB 474 GGDWRNWRSELYKYVKVTEPLGVAPTAKRRVQREERAA-AIGALFLGLGAGSTMG 532

QY 541 AASVTLTVQAROLLSGIVOOQSNLLRAIEAQHLLQLTVWGIKQLOARILAVERYLKDOQ 600
 DB 533 AASVTLTVQARLLSGIVOOQNNLLRAIEAQHLLQLTVWGIKQLOARVQAVERYLKDOQ 592
 QY 601 LLGIMGCSGKLICTTTPWNASWSKNKTLEAIVNNMTWQWDEIDNTYLSIYSLIEESP 660
 DB 593 LLGFWGCSGKLICTTTPWNASWS-NKSLDDIWNNTWQWDEIDNTYLSIYSLLEKSQ 651
 QY 661 IQQKNEQBELLEDKQANLWNPFDISNWLWYIKIFIMIVGGLVGLRIVFVVLIVNRVRQ 720
 DB 652 TQQKNEQBELLEDKQASLWNPFDITNWLWYIKIFIMIVGGLVGLRIVFAVLIVNRVRQ 711
 QY 721 GYSPISFOTLPAPGPDPEETEEEGGDRDRSGLLVGDGFLTLIWDLSLCLFSYHR 780
 DB 712 GYSPISLQTRPPVPRGPDPEETEEEGGDRDRSGRLVHGFLAIIVDDURSLFLSYHH 771
 QY 781 LRDLIIIVTRIVELLCRRGWEILKYWNLLQYWSQELKNSAVSLFNATAIAAEGTDRVI 840
 DB 772 LRDLIIIAIRIVELLCRRGWEILKYWNLLQYWSQELKSSAVSLNATAIAAEGTDRVI 831
 QY 841 QVLQVRGALLHIPTIRQGLERALL 866
 DB 832 EVLQRAIRAILHIPTIRQGLERALL 857
 RESULT 4
 ENV_HV1MN STANDARD; PRT; 856 AA.
 ID ENV_HV1MN
 AC P05877;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Envelope polyprotein GP160 precursor (Contains: Exterior membrane
 glycoprotein (GP120); Transmembrane glycoprotein (GP41)).
 GN Name=Env;
 OS Human immunodeficiency virus type 1 (MN isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11696;
 [1]
 SEQUENCE FROM N.A.
 RX MEDLINE=88219542; PubMed=3369091;
 RA Gurgo C., Guo H.-G., Franchini G., Aldovini A., Collalti E.,
 Farrell K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;
 RA "Envelope sequences of two new United States HIV-1 isolates.";
 RL Virology 164:531-536(1988).
 CC -!- MISCELLANEOUS: The MN isolate was taken from a pediatric AIDS
 patient in 1984.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; M17449; AAA44857.1; -.
 DR PDB; 1ACV; X-ray; P=306-329.
 DR PDB; 1F58; X-ray; P=306-328.
 DR PDB; 1NZ; NMR; A=309-324.
 DR PDB; 1NJ0; NMR; A=309-324.
 DR HIV; M17449; ENVSMN.
 DR InterPro; IPR000328; Env_GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KW 3D-structure; AIDS; Coat protein; Glycoprotein; Polyprotein; Signal;
 Transmembrane.
 FT SIGNAL 1 29
 FT CHAIN 30 513 Exterior membrane glycoprotein.
 FT CHAIN 514 856 Transmembrane glycoprotein.
 FT DISULFID 53 73 By similarity.
 FT DISULFID 118 210 By similarity.

FT	DISULFID	125	201	By similarity.	
FT	DISULFID	130	162	By similarity.	
FT	DISULFID	223	252	By similarity.	
FT	DISULFID	233	244	By similarity.	
FT	DISULFID	301	335	By similarity.	
FT	DISULFID	381	445	By similarity.	
FT	DISULFID	388	418	By similarity.	
FT	CARBOHYD	87	87	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	129	129	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	135	135	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	140	140	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	141	141	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	146	146	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	161	161	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	165	165	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	191	191	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	202	202	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	246	246	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	267	267	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	281	281	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	294	294	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	300	300	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	336	336	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	343	343	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	359	359	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	365	365	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	395	395	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	401	401	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	405	405	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	406	406	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	413	413	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	448	448	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	465	465	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	612	612	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	617	617	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	626	626	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	638	638	N-linked (GLCNAC..)	(Potential).
FT	CARBOHYD	816	816	N-linked (GLCNAC..)	(Potential).
SQ	SEQUENCE	856 AA;	97140 MW;	D197D805940BE732	CRC64;

Query Match 86.0%; Score 3970; DB 1; Length 856;
Best Local Similarity 86.7%; Pred. No. 2.7e-285;
Matches 751; Conservative 36; Mismatches 69; Indels 10; Gaps 6;

QY	1	MRVKGIRRYQHWWGWTMLGLMLTCSATEKLWVTVYGVVPVWKEATTLFCASDAKAY	60
Db	1	MRVKGIRRYQHWWGWTMLGLMLTCSATEKLWVTVYGVVPVWKEATTLFCASDAKAY	60
QY	61	DTBAHNVWATHACVPTDPNPQVELVNVTFNFMNKNMVEQMHEDIISLWQSLKPCVK	120
Db	61	DTEVHNWATQACVPTDPNPQVELVNVTFNFMNKNMVEQMHEDIISLWQSLKPCVK	120
QY	121	LTPLCVTLACTDLRNTTNTNSTDNNSSEGTIKGEMKNCNFNATSGDKMKQKEYAL	180
Db	121	LTPLCVTLNCTDLRNTTNTNSTANNNSSEGTIKGEMKNCNFNTTSIRDMMQKEYAL	180
QY	181	LYKLDIEPIDNMTSVRLTSCNTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG	240
Db	181	LYKLDIVSINDSTSVRLTSCNTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG	240
QY	241	KGSKCNVSTVQCTHGIRPVVSTQLLNGSLAEVEVIRSENFTNNAKTIIVQUREPVKIN	300
Db	241	KGSKCNVSTVQCTHGIRPVVSTQLLNGSLAEVEVIRSENFTNNAKTIIVHLNESVQIN	300
QY	301	CSPNNNTRKSIIPMGGRAPYTTGIIIGDROACHNISKNTWNTNALKQVVKLGEQFNKT	360
Db	301	CTRPNTNKKRIHIGRGRAPYTTKNTIGTIRQACHNISRANKWNTLRQIVSKLGEQFNK	360
QY	361	KIVFTNSSGGDPEIVTHSFNCAGEFFVCNTTOLFDSIWNSENCTWNTIRGLNNTGNDTI	420
Db	361	TIVFNSSGGDPEIVHMSFNCGGEFFVCNTSPFNSTWNG-NNTWN-----NTTGSNNNI	414
QY	421	TLPCRIRKQIINRWQEVGKAMYPPIKGNISCSNITGLLLTRDGGKDDNSRDGNETFRFG	480

Db	415	TLQCKIKQIINRWQEVGKAMYPPIEGQIRCSNITGLLLTRDGGKDDTND-TEIFRFG	473
QY	481	GGDMRDWSELYKYKVKIIEPLGVAPTAKERVVOREBVRVGLGAMFFGELGAAGSTWG	540
Db	474	GGDMRDWSELYKYKVKVITIEPLGVAPTAKERVVOREKRA-AIGALFLGFLGAAGSTWG	532
QY	541	AASVTITVQARQLLSGIVQQSNLLRAIEAQOHLQLTWGIKQLQARILAVERYLKDOQ	600
Db	533	AASVTITVQARQLLSGIVQQSNLLRAIEAQOHLQLTWGIKQLQARVLAVERYLKDOQ	592
QY	601	LLGWCSCGKLICTTTPVWNASWKNKTLEAIWNMTWQWDEIDNTYSLIYSLIEESP	660
Db	593	LLGFWCSCGKLICTTTPVWNASWS-NKSLDDIWNMTWQWEREIDNTYSLIYSLIEKSQ	651
QY	661	IQQKNEQELLELDKWNANWNPFDISNWLWYIKIFIMIVGGLVGLRIVFVWLSIVNRVQ	720
Db	652	TOQKNEQELLELDKWNANWNPFDITNWLWYIKIFIMIVGGLVGLRIVFAVLSIVNRVQ	711
QY	721	GYSPISFOTRLPAPGPDPRPEIEBEGGDRDRSGLLVGDFLTLIWDLRSICLFSYHR	780
Db	712	GYSPISLQTRPVPGRPDPRPEIEBEGGDRDTSGLVHGFLAIWVWDLRSILFSYHH	771
QY	781	LRDLLLIVTRIVELIGRRGWEILKYWNLLQYWSQELKNSAVSLFNATAIAVAEGTDRVI	840
Db	772	-RDLLLIAARIVELLIGRRGWEVLKYWNLLQYWSQELKSSAVSLNATAIAVAEGTDRVI	830
QY	841	QVLQVRGALLHIPTRIOGLERALL	866
Db	831	EVLQAGRAILHIPTRIOGLERALL	856

RESULT 5
QYUL5 PRELIMINARY; PRT; 857 AA.
ID QYUL5; AC QYUL5; DT 01-MAY-1999 (TREMELrel. 10, Created)
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98362112; PubMed=9696803;
RA Park E.J., Vujcic L.K., Anand R., Theodore T.S., Quinnan G.V. Jr.;
RT "Mutations in both gp120 and gp41 are responsible for the broad
RT neutralization resistance of variant human immunodeficiency virus type
RT 1 MN to antibodies directed at V3 and non-V3 epitopes.";
RL J. Virol. 72:7099-7107(1998).
DR EMBL; AF075722; AAC31819.1; -.
DR HSSP; P05877; INQ0.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;
KW Transmembrane.
SQ SEQUENCE 857 AA; 97095 MW; 922D362E10C60DA8 CRC64;

Query Match 86.0%; Score 3969.5; DB 2; Length 857;
Best Local Similarity 86.4%; Pred. No. 2.9e-285;
Matches 748; Conservative 41; Mismatches 68; Indels 9; Gaps 5;

QY	1	MRVKGIRRYQHWWGWTMLGLMLTCSATEKLWVTVYGVVPVWKEATTLFCASDAKAY	60
Db	1	MRVKGIRRYQHWWGWTMLGLMLTCSATEKLWVTVYGVVPVWKEATTLFCASDAKAY	60

Db 712 GYSPLSQTQTPVPVPRGDRPEGIEBEGGERDRDTSGLRVHGFALAIIVWDLRLSFLLSYHH 771
 Qy 781 LRDLIIIVTRIVELLGRGWEILKYWNLLQYWSQELKNSAVLNFATAIAVAEGTDRVI 840
 Db 772 LRDLIIAARIVELLGRGWEILKYWNLLQYWSQELKNSAVLNFATAIAVAEGTDRVI 831
 Qy 841 QVLQVRGRALLHIPTRIQGLERALL 866
 Db 832 EVLQVRAGRALLHIPTRIQGLERALL 857

RESULT 7
 O90296 PRELIMINARY; PRT; 857 AA.
 AC O90296;
 DT 01-NOV-1998 (T-EMBLrel. 08, Created)
 DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)
 DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
 DE Envelope glycoprotein.
 GN Name=env;
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 [1]
 RN SEQUENCE FROM N.A.
 RA MEDLINE=98362112; PubMed=9696803;
 RX Park E.J., Vujcic L.K., Anand R., Theodore T.S., Quinlan G.V. Jr.;
 RT "Mutations in both gp120 and gp41 are responsible for the broad
 neutralization resistance of variant human immunodeficiency virus type
 RT 1 MN to antibodies directed at V3 and non-V3 epitopes."
 RL J. Virol. 72:7099-7107(1998).
 DR EMBL; AF075721; AAC31818.1; -.
 DR HSP; P05877; INJ0.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; F:structural molecule activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000328; Env GP41.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polypotein;
 KW Transmembrane.
 SQ SEQUENCE 857 AA; 97064 MW; E0D8599CC50C2F47 CRC64;

Query Match 85.4%; Score 3941.5; DB 2; Length 857;
 Best Local Similarity 85.7%; Pred. No. 3.5e-283;
 Matches 742; Conservative 46; Mismatches 69; Indels 9; Gaps 5;

Qy 1 MRVKGIRRNQHWGWTMLGLLMICSAATEKLWTVTVYGVVYVKEATTLFCASDAKAY 60
 Db 1 MRVKGIRRNQHWGWTMLGLLMICSAATEKLWTVTVYGVVYVKEATTLFCASDAKAY 60
 Qy 61 DTAHNWATHACVPTDPNPOVELNVNFTFNWKNMVEQHEHDIISLDQSLKPCVK 120
 Db 61 DTAHNWATHACVPTDPNPOVELNVNFTFNWKNMVEQHEHDIISLDQSLKPCVK 120
 Qy 121 LTPCLVTLNCTDLRNTNTNNSSEGTIGKGMKNCFSNATSGDKMKEAYAL 180
 Db 121 LTPCLVTLNCTDLRNTNTNNSSEGTIGKGMKNCFSNATSGDKMKEAYAL 180
 Qy 181 LYKLDIEPIDNDNTSYRLISNTSVITQACPKISFEPIPIHYCAPAGFAILKCNCKKFSG 240
 Db 181 LYKLDIEPIDNDNTSYRLISNTSVITQACPKISFEPIPIHYCAPAGFAILKCNCKKFSG 240
 Qy 241 KGSCKNVSTVQCHGIRPVVSTOLLNGSLAEBEVIRSENFNTNNAKTIIVQREPVKIN 300
 Db 241 KGSCKNVSTVQCHGIRPVVSTOLLNGSLAEBEVIRSENFNTNNAKTIIVHLSKESVQIN 300
 Qy 301 CSRPNNTKRSIPMGPGRAFTTQGIIGDQIAHCNISKNTNWNALKQVVEKLGEOFNKT 360
 Db 301 CTRPNWNRKRIHIGPGRAFTTQGIIGDQIAHCNISKNTNWNALKQVVEKLGEOFNKT 360

Qy 361 KIVFTNSSGDDEIIVTHSFNCAGEFFYCNNTQLPDSIWNSENGTWNITRGLNNTGNDTI 420
 Db 361 TIVFNPSSGDDEIIVHMFNCAGEFFYCNNTQLPDSIWNSENGTWNITRGLNNTGNDTI 414
 Qy 421 TLPCRKQIILNWRQVGVKAMVAPPIKGNISCSNITGILLTRDGGKDDNSRDGNETFRG 480
 Db 415 TIQCKVKQIILNWRQVGVKAMVAPPIEGQIRCSNITGILLTRDGGEDTDND-TIIFRPG 473
 Qy 481 GGDMDNRSELYKYKVKVIEPLGVAPTAKKRVVQREARVGLGAMFFGFLGAAGSTWG 540
 Db 474 GGDMDNRSELYKYKVKVIEPLGVAPTAKKRVVQREKRA-AIGALFLGLGAAGSTWG 532
 Qy 541 AASVTLTVQAROLLGSGIVQOQNLLRAIEAQHLLQLTWGIKQIQARILAVERYLKQOQ 600
 Db 533 AASVTLTVQAROLLGSGIVQOQNLLRAIEAQHLLQLTWGIKQIQARVQAVERYLKQOQ 592
 Qy 601 LIGWGCCKLCTTTPVWNASWSKNKTLKLNWNTWQWDEIDNTYSLIYSLIESP 660
 Db 593 LIGWGCCKLCTTTPVWNASWS-NKSLDDIWNWNTWQWDEIDNTYSLIYSLIESQ 651
 Qy 661 IQQEKNEQELLEDKWANLWNFDISNMLYIKIFIMIVGGIVGLRIVFVVLISIVNRVQ 720
 Db 652 TQQEKNEQELLEDKWESLWNFDISNMLYIKIFIMIVGGIVGLRIVFAVLISIVNRVQ 711
 Qy 721 GYSPLSQTQTPVPVPRGDRPEGIEBEGGERDRDTSGLRVHGFALAIIVWDLRLSFLSYH 780
 Db 712 GYSPLSQTQTPVPVPRGDRPEGIEBEGGERDRDTSGLRVHGFALAIIVWDLRLSFLSYH 771
 Qy 781 LRDLIIIVTRIVELLGRGWEILKYWNLLQYWSQELKNSAVLNFATAIAVAEGTDRVI 840
 Db 772 LRDLIIAARIVELLGRGWEILKYWNLLQYWSQELKNSAVLNFATAIAVAEGTDRVI 831
 Qy 841 QVLQVRGRALLHIPTRIQGLERALL 866
 Db 832 EVLQVRAGRALLHIPTRIQGLERALL 857

RESULT 8
 Q7SVL6 PRELIMINARY; PRT; 853 AA.
 AC Q7SVL6;
 DT 01-OCT-2003 (T-EMBLrel. 25, Created)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 DE Envelope glycoprotein.
 GN Name=env;
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=22628496; PubMed=12743293;
 RX DOI=10.1128/JVI.77.11.6359-6366.2003;
 RA Robbins K.E., Lemey P., Pybus O.G., Jaffe H.W., Youngpairoj A.S.,
 RA Brown T.M., Salemi M., Vandamme A.M., Kalish M.L.;
 RT "U.S. Human immunodeficiency virus type 1 epidemic: date of origin,
 RT population history, and characterization of early strains."
 RL J. Virol. 77:6359-6366(2003).
 DR EMBL; AY247219; AAP37147.1; -.
 DR HSP; P20871; ICE4.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; F:structural molecule activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000328; Env GP41.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polypotein;
 KW Transmembrane.
 SQ SEQUENCE 853 AA; 96545 MW; E3B2830A1261E237 CRC64;

Query Match 85.2%; Score 3931.5; DB 2; Length 853;

Best Local Similarity 85.1%; Pred. No. 1.9e-282;
Matches 739; Conservative 53; Mismatches 59; Indels 17; Gaps 6;

QY 1 MRVKGIRRYQHWWGWTMLLGLLMICSAATEKLWVTYVYGVVPVWKEATTLFPCASDAKAY 60
DB 1 MRVKGIRKYNQHLWRGWTMLLGLLMICSAATEKLWVTYVYGVVPVWKEATTLFPCASDAKAY 60

QY 61 DTAHNWATHACVPTDNPQEVLELVNVTENFNMKNMVEQMHEDIISLDOSLKPVCVK 120
DB 61 DTEVHNWATHACVPTDNPQEVLELVNVTENFNMKNMVEQMHEDIISLDOSLKPVCVK 120

QY 121 LTPCLVTLNCTDLRNTNTNNSSEGTIKGEMKNCNFNATSIIGDKMQKEYAL 180
DB 121 LTPCLVTLNCTDLRNTNTNNSSEGTIKGEMKNCNFNATSIIGDKMQKEYAL 173

QY 181 LYKLDIEPDND--NTSYRLISNTSVITQACPKISFEPIPIHYCAPAGFAILKCNDDKKF 238
DB 174 FYKLDVVPIDNDNTNTSYRLISNTSVITQACPKISFEPIPIHYCAPAGFAILKCNDDKKF 233

QY 239 SGKSGCKNVSTVQCTHGIRPVSTQLLNGSLAEVEVIRSENFTNNAKTIIVQLREPVK 298
DB 234 SGKGPCTNVSTVQCTHGIRPVSTQLLNGSLAEVEVIRSDNFTNNAKTIIVQLKEPVE 293

QY 299 INCSRPNNTKSIIPMGPGRAFYTTQGIIGDIRQAHNCNISKTNWTKALQVVEKLGQFN 358
DB 294 INCTRPNNTRKSIHIGPGRAFYTTGEGIIIGDIRQAHNCNLSRAKWDNTLQIVVKLRQFG 353

QY 359 KTKIVFTNSGGDPEIIVTHSFNCAGFEFFYCNVTQTFDSIWNSENGTWNITRGLNTRGRND 418
DB 354 NKTIVFNQSSGGDPEIIVTHSFNCAGFEFFYCNVTQTFDSIWNSENGTWNITRGLNTRGRND 410

QY 419 TITLPCRIKQIINRWQEVGKAMYPPIKGNISCSNITGLLLTRDGGKDNDRDNETFR 478
DB 411 -ITLPCRIKQIINRWQEVGKAMYPPIKGNISCSNITGLLLTRDGGKDNDRDNETFR 466

QY 479 PGCGDMRDNRSELYKYKVKIEPLGVAPTAKRRVVQREKRAVGLGAMFPGFLGAAGST 538
DB 467 PGCGDMRDNRSELYKYKVKIEPLGVAPTAKRRVVQREKRAVGLGAMFPGFLGAAGST 526

QY 539 MGAASVTLTVQARQLLSGIVQOQSNLLRAEAQOHLQLTWGIIKQIQAARILAVERYLKD 598
DB 527 MGAASVTLTVQARQLLSGIVQOQSNLLRAEAQOHLQLTWGIIKQIQAARILAVERYLKD 596

QY 599 QQLLGWCGSGKLICTTTPVWNASWSKNKTLEAIWNNMTWMDKEIDNYSLSIYLIEE 658
DB 587 QQLLGWCGSGKLICTTTPVWNASWS--NKSLDKIWNMTWMEWEREINNTYSLSIYLIEE 645

QY 659 SPIQEKNEQELLELDKWLWNLWFDISNWLWIKIFIMIVGLVGLRIVFVVLISVNRV 718
DB 646 SONQEKNELELELDKWLWNLWFDITKWLWIKIFIMIVGLVGLRIVFVVLISVNRV 705

QY 719 ROGYSPLSFOTRLPAPRGPDRPEIEEGGDRDRDRSGLLVGDFLTLIWDLSLCLFSY 778
DB 706 ROGYSPLSFOTRLPAPRGPDRPGIEEGGKDRDRSGLLVGDFLTLIWDLSLCLFSY 765

QY 779 HRLRDLILLIVTRIVELLGREGWELKYWNLLQYWSQELKNSAVSLFNATAIAVAGTDR 838
DB 766 HRLRDLILLIVTRIVELLGREGWELKYWNLLQYWSQELKNSAVSLFNATAIAVAGTDR 825

QY 839 VIQVQRVGRALHPIPRIOGLERALL 866
DB 826 VIEWLQRAYRAILHPIPRIOGLERALL 853

RESULT 9
Q7SVL4
ID Q7SVL4 PRELIMINARY; PRT; 850 AA.
AC Q7SVL4;
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 25, Last annotation update)
DE Envelope glycoprotein.
GN Name=env;

OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=22628496; PubMed=12743293;
RX DOI=10.1128/JVI.77.11.6359-6366.2003;
RA Robbins K.E., Lemey P., Pybus O.G., Jaffe H.W., Youngpairaj A.S.,
Brown T.M., Salemi M., Vandamme A.M., Kalish M.L.;
RT "U.S. Human immunodeficiency virus type 1 epidemic: date of origin,
population history, and characterization of early strains."
J. Virol. 77:6359-6366(2003).
RL EMBL; AY247221; AAP37149.1; -;
DR HSSP; P20871; ICE4.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;
Transmembrane.
SQ SEQUENCE 850 AA; 96583 MW; 62ED5F6AB033D20B CRC64;

Query Match 85.1%; Score 3928; DB 2; Length 850;
Best Local Similarity 85.3%; Pred. No. 3.5e-282;
Matches 740; Conservative 45; Mismatches 63; Indels 20; Gaps 5;

QY 1 MRVKGIRRYQHWWGWTMLLGLLMICSAATEKLWVTYVYGVVPVWKEATTLFPCASDAKAY 60
DB 1 MRVKGIRKYNQHLWRGWTMLLGLLMICSAATEKLWVTYVYGVVPVWKEATTLFPCASDAKAY 60

QY 61 DTAHNWATHACVPTDNPQEVLELVNVTENFNMKNMVEQMHEDIISLDOSLKPVCVK 120
DB 61 DTEVHNWATHACVPTDNPQEVLELVNVTENFNMKNMVEQMHEDIISLDOSLKPVCVK 120

QY 121 LTPCLVTLNCTDLRNTNTNNSSEGTIKGEMKNCNFNATSIIGDKMQKEYAL 180
DB 121 LTPCLVTLNCTDLRNTNTNNSSEGTIKGEMKNCNFNATSIIGDKMQKEYAL 173

QY 181 LYKLDIEPDND--TSYRLISNTSVITQACPKISFEPIPIHYCAPAGFAILKCNDDKKF 238
DB 174 FYKLDVVPIDNDNTNTSYRLISNTSVITQACPKISFEPIPIHYCAPAGFAILKCNDDKKF 233

QY 239 SGKSGCKNVSTVQCTHGIRPVSTQLLNGSLAEVEVIRSENFTNNAKTIIVQLREPVK 298
DB 234 NGTGPCTNVSTVQCTHGIRPVSTQLLNGSLAEVEVIRSDNFTNNAKTIIVQLKEPVE 293

QY 299 INCSRPNNTKSIIPMGPGRAFYTTQGIIGDIRQAHNCNISKTNWTKALQVVEKLGQFN 358
DB 294 INCTRPNNTRKSIHIGPGRAFYTTGEGIIIGDIRQAHNCNLSRAKWDNTLQIVVKLRQFG 353

QY 359 KTKIVFTNSGGDPEIIVTHSFNCAGFEFFYCNVTQTFDSIWNSENGTWNITRGLNTRGRND 418
DB 354 NKTIVFNQSSGGDPEIIVTHSFNCAGFEFFYCNVTQTFDSIWNSENGTWNITRGLNTRGRND 406

QY 419 TITLPCRIKQIINRWQEVGKAMYPPIKGNISCSNITGLLLTRDGGKDNDRDNETFR 478
DB 407 TITLPCRIKQIINRWQEVGKAMYPPIKGNISCSNITGLLLTRDGGKDNDRDNETFR 463

QY 479 PGCGDMRDNRSELYKYKVKIEPLGVAPTAKRRVVQREKRAVGLGAMFPGFLGAAGST 538
DB 464 PGCGDMRDNRSELYKYKVKIEPLGVAPTAKRRVVQREKRAVGLGAMFPGFLGAAGST 523

QY 539 MGAASVTLTVQARQLLSGIVQOQSNLLRAEAQOHLQLTWGIIKQIQAARILAVERYLKD 598
DB 524 MGAASVTLTVQARQLLSGIVQOQSNLLRAEAQOHLQLTWGIIKQIQAARILAVERYLKD 583

QY 599 QQLLGWCGSGKLICTTTPVWNASWSKNKTLEAIWNNMTWMDKEIDNYSLSIYLIEE 658
DB 584 QQLLGWCGSGKLICTTTPVWNASWS--NKSLDKIWNMTWMEWEREINNTYSLSIYLIEE 642

DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;
KW Transmembrane.
SQ SEQUENCE 860 AA; 97762 MW; A7E185F7BD421590 CRC64;

Query Match 84.6%; Score 3905; DB 2; Length 860;
Best Local Similarity 84.2%; Pred. No. 1.8e-280;
Matches 734; Conservative 54; Mismatches 66; Indels 18; Gaps 7;

QY 1 MRVKGIRRYQHWWGCTMLGLLMTICSAATEKLWVTYGVVPWKATTTLFCASDAKAY 60
DB 1 MRVKGIRKYNQHWWGCTMLGLLMTICSAAGKLVVTYGVVPWKATTTLFCASDAKAY 60
QY 61 DTEAHNVWATHACVPTDNPQVEVLNVTFENFMNKNMVEQMHEDIISLWDSQKPCVK 120
DB 61 DTEIHNWATHACVPTDNPQVEALENVTFENFMNKNMVEQMHEDIISLWDSQKPCVK 120
QY 121 LTPCLVTLNCTDLRNTNTNNSNSSEGTIKGEMKNCSEFNATSIIGDKMKEAYAL 180
DB 121 LTPCLVTLNCTDLRNTNTNNSNSSEGTIKGEMKNCSEFNATSIIGDKMKEAYAL 173
QY 181 LYKLDIEPIDND--NTSYRLISNTSVITQACPKISPEPIHYCAPAGFAILKCNKDKF 238
DB 174 FYRLDVPIDNDNTSYRLISNTSVITQACPKVSEPIHYCAPAGFAILKCNKDKF 233
QY 239 SGKSGKNVSTVQCTHGIKRVWSTQILLNGSLAEVEVIRSENFTNNAKTIIVQLREPVK 298
DB 234 NGTGPCNTVSTVQCTHGIKRVISTQILLNGSLAEVEVIRSDNFMNNAKTIIVQLNESVE 293
QY 299 INCSRPNNTKSIIPMGPGRAFTYTGIIIGDIQAHNCISKTNTWTKALQVVEKLGEPFN 358
DB 294 INCTRPNNTRKSIHIGPGRAFTYTGIIIGDIQAHNCISKTNTWTKALQVVEKLGEPFN 353
QY 359 KTIKIVTNSGGDPEIVTHSFNCAGFEFFYNTTQDFSIWNSB---NGTW-NITRGLNNT 414
DB 354 NKTIVFNQSGGDPEIVTHSFNCAGFEFFYCDTQDFNSTWNTDNLNSTWNTDNGKSNNT 413
QY 415 GRNDTITLPCRIKQIINRWQEVGKAMYAPPIKGNISCSNITGLLITRDCGKDDNSRDGN 474
DB 414 --EDTITLPCRIKQIINRWQEVGKAMYAPPIKQIRCSNITGLLITRDCG--NNKSETT 469
QY 475 ETRPGGDMRDNWRSELYKYKVKIIEPLGVAFTKAKRRVVRQREAVGLGAMPFGFLGA 534
DB 470 EIERPGGDMRDNWRSELYKYKVKIIEPLGVAFTKAKRRVVRQREAVGLGAMPFGFLGA 529
QY 535 AGSTMGAASVTLTVQARQLLSGIVQOOSNLLRAIEAQHLLQITVWGIKQLOARILAVR 594
DB 530 AGSTMGAASMTLTVQARQLLSGIVQOQNLLRAIEAQHLLQITVWGIKQLOARILAVR 589
QY 595 YLKDQQLLGIWCGSGKLICTTTPVWNASWKNKTLRAIMNMTWMDKEIDNYSILYS 654
DB 590 YLKDQQLLGIWCGSGKLICTTTPVWNASW--NKSCLKIMNMTWMEWEINNTSYLIY 648
QY 655 LIESPQKNEQKQELLELDKMANLWNPDI SNLWYIKIFIMVGLGLRIVFVLSI 714
DB 649 LIESQKQKNEQKQELLELDKMANLWNPDI TWLWYIKIFIMVGLGLRIVFVLSI 708
QY 715 VNRVQGSPLSPTQLPAPRGPDRPEIEEGGDRDRSGLLVDGFLTLIVWDLRLSLC 774
DB 709 VNRVQGSPLSPTQLPAPRGPDRPEIEEGGDRDRSGRLVDGFLFALFVWDLRLSLC 768
QY 775 LFSYHRLDLLLIVTRIVELLGRGWEILKYWNLLQYMSQELKNSAVSLFNATAVAE 834
DB 769 LFSYHRLDLLLIVTRIVELLGRGWEILKYWNLLQYMSQELKNSAVSLFNATAVAE 828
QY 835 GTDRVITQVLRGALLHIPTIRIQGLERALL 866
DB 829 GTDRITVLRVRAIHIPTIRIQGLERALL 860

RESULT 12

092763 PRELIMINARY; PRT; 863 AA.
AC 092763;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN Name-env.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98178716; PubMed=9519894;
RA McCutchan F.E.; Sanders-Buell E.; Salminen M.O.; Carr J.K.,
RT "Diversity of the human immunodeficiency virus type 1 envelope
glycoprotein in San Francisco Men's Health Study participants.";
RL AIDS Res. Hum. Retroviruses 14:329-337(1998).
DR EMBL; AF025755; AAC40592.1; -;
DR HSSP; P04578; 1DLB.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;
KW Transmembrane.
SQ SEQUENCE 863 AA; 97863 MW; 13A9530DBBAEF4C CRC64;

Query Match 84.4%; Score 3896.5; DB 2; Length 863;
Best Local Similarity 84.3%; Pred. No. 7.7e-280;
Matches 734; Conservative 49; Mismatches 75; Indels 13; Gaps 6;

QY 1 MRVKGIRRYQHWWGCTMLGLLMTICSAATEKLWVTYGVVPWKATTTLFCASDAKAY 60
DB 1 MRVKGIRKYNQHWWGCTMLGLLMTICSAAEKLWVTYGVVPWKATTTLFCASDAKAY 60
QY 61 DTEAHNVWATHACVPTDNPQVEVLNVTFENFMNKNMVEQMHEDIISLWDSQKPCVK 120
DB 61 DTEIHNWATHACVPTDNPQVEVLNVTFENFMNKNMVEQMHEDIISLWDSQKPCVK 120
QY 121 LTPCLVTLNCTDLRNTNTNNSNSSEGTIKGEMKNCSEFNATSIIGDKMKEAYAL 180
DB 121 LTPCLVTLNCTDLRNTNT--TANTNSSGGAMEEGDKNCSEFNITTSIRNMQRERVL 177
QY 181 LYKLDIEPIDNDNTSVITQACPKISPEPIHYCAPAGFAILKCNKDKFSG 240
DB 178 FYKLDIVPIDNDNTSVITQACPKVSEPIHYCAPAGFAILKCNKDKFSG 237
QY 241 KGSKNVSTVQCTHGIKRVWSTQILLNGSLAEVEVIRSENFTNNAKTIIVQLREPVKIN 300
DB 238 TGPCENVSTVQCTHGIKRVWSTQILLNGSLAEVEVIRSENFTNNAKTIIVQLNASVEIN 297
QY 301 CSRPNNTKSIIPMGPGRAFTYTGIIIGDIQAHNCISKTNTWTKALQVVEKLGEPFN-K 359
DB 298 CTRPNNTKSIHIGPGRAFTYTGIIIGDIQAHNCISKTNTWTKALQVVEKLGEPFNK 357
QY 360 TKIVFTNSGGDPEIVTHSFNCAGFEFFYNTTQDFSIWNSBNGTWNITRGLNNTGRNDT 419
DB 358 TTVTFNQSGGDPEVTVTHSFNCAGFEFFYCNSTKLFNSTWKF--NNTWNTSNT--DT 413
QY 420 ITPPCRIKQIINRWQEVGKAMYAPPIKGNISCSNITGLLITRDCGKDDNSRDGN---E 475
DB 414 ITPPCRIKQIINRWQEVGKAMYAPPIKGNISCSNITGLLITRDCGKDDNSRDGNETE 473
QY 476 TFRPGGDMRDNWRSELYKYKVKIIEPLGVAFTKAKRRVVRQREAVGLGAMPFGFLGAA 535
DB 474 IFRPGGDMRDNWRSELYKYKVKIIEPLGVAFTKAKRRVVRQREAVGLGAMPFGFLGAA 533

```
Qy 536 GSTGGAASVLTVOARQLLSGIVQOQSNLLRAEAQOHLQLTVMGKIQOARILAVERY 595
Db 534 GSTGGAASITLTVOARQLLSGIVQOQSNLLRAEAQOHLQLTVMGKIQOARVLAVERY 593
Qy 596 LKQOQLLGIWGCCKLICTTTPVNASWSKNKLTLEATWNNMTWMDKEIDNTSLIYSL 655
Db 594 LVPQQLLGIWGCCKLICTTAVPNASWS-NKSPKEIWNNTWWEKEINNTRLIYTL 652
Qy 656 IEESPIQOEKNEBELLELDKAWNLNNWFDISNLLWYIKIFIMIVGGLVGLRIVFVLSIV 715
Db 653 IEESQOEKNEBELLELDKAWNLNNWFDITNLLWYIRIFIMIVGGLVGLRIVFVLSIV 712
Qy 716 NRVQGYSPLSFOTRLPAPRDPREPBEIEGGDRDRSGLLVGDLTLIWDLSLCL 775
Db 713 NRVQGYSPLSLQTRPPAPRDPREPBEIEGGDRDRSGLLVGDLTLIWDLSLCL 772
Qy 776 FSHRLRDLALLIVTRIVELLGRGWELTKYWNLLQYWSQELKNSAVLSFNATAIAVAG 835
Db 773 FSHRLRDLALLIVTRIVELLGRGWELTKYWNLLQYWSQELKNSAVLSFNATAIAVAG 832
Qy 836 TDRVIOVLQVRGRALLHIPTRIQGLERALL 866
Db 833 TDRIVVQRTFRALLHIPTRIQGLERALL 863

RESULT 13
Q7SVL5
ID Q7SVL5 PRELIMINARY; PRT; 852 AA.
AC Q7SVL5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Envelope glycoprotein.
GN Name:env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX DOI=10.1128/JVI.77.11.6359-6366.2003;
RA Robbins K.E., Lemey P., Pybus O.G., Jaffe H.W., Youngpairoj A.S.,
  Brown T.M., Salemi M., Vandamme A.M., Kalish M.L.;
  "U.S. Human immunodeficiency virus type 1 epidemic: date of origin,
  RT population history, and characterization of early strains.";
  J. Virol. 77:6359-6366(2003).
DR EMBL; AY247220; AAP37148.1; -.
DR HSP; P04578; IDLB.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;
  Transmembrane.
SQ SEQUENCE 852 AA; 96680 MW; F8D811028025D998 CRC64;

Query Match 84.2%; Score 3886; DB 2; Length 852;
Best Local Similarity 84.6%; Pred. No. 4.5e-279;
Matches 733; Conservative 50; Mismatches 69; Indels 14; Gaps 5;

Qy 1 MRVKGIRRNQHWGWTMLGLMLTCSATEKLWTVTVYGVVPVKEATTLFCASDAKAY 60
Db 1 MRVKEIGKTYLWRGWTMLGLMLTCSAAEQULWTVYGVVPVKEATTLFCASDAKAY 60
Qy 61 DTEAHNVWATHACVPTDPNPQEVVELNVNTEFNWKNMVEQWHEDIISLWQSLKPCVK 120
Db 61 DTEHNVWATHACVPTDPNPQEVGLENVTEFNWKNMVEQWHEDIISLWQSLKPCVK 120
Qy 121 LTPCLVTNCTDLGNTNT- TSSGGGMEGEIKNCSFNITTSIRDKVQKEYAL 173
```

```
Db 121 LTPCLVTNCTDLGNTNT- TSSGGGMEGEIKNCSFNITTSIRDKVQKEYAL 173
Qy 181 LYKLDIEPIDDNNTSVRLISCNSTSVITQACPKISPEPIPIHYCAPAGFALLKCNKKFSG 240
Db 174 LYKLDIVPIDDNNTSVRLISCNSTSVITQACPKVSFPEPIPIHYCAPAGFALLKCRDKKFG 233
Qy 241 KGSCKNVSTVQCTHGIRPVVSTQLLNGSLABEEVVIRSEFTNNAKTIIVQLREPVKIN 300
Db 234 KGPCNTVSTVQCTHGIRPVVSTQLLNGSLABEEVVIRSDNFTDNAKTIIVQLKESVEIN 293
Qy 301 CSRPNNTKSGIPMGPRAFYTTGQIIGDIRQAHCMISKTNTNALKQVVEKLGGEFKNKT 360
Db 294 CTRPNNTKSGIHIGPRAFYTTGDIVGDIRQAHCMISKAKWNTTLKQIVTKLREQFGNR 353
Qy 361 KIVFTNSSGDEIIVTHSFNCAGEFFYCNLTQFDSIWNSENGTWNITGLNNTGRNDTI 420
Db 354 TIVFNOSGDEIIVHMSFCGGEFFYCNLTQFNSWNA-SSTWNTDTSNNNT--EGTI 410
Qy 421 TLPCKRIKQIINRWQEVGKAMYAPPIKGNISCSNITGLLLTRDGGKDDRSRDNCTFRPG 480
Db 411 TLPCKRIKQIINRWQEVGKAMYAPPIRGQIRCSNITGLLLTRDGGSNEND--TEIFRPG 467
Qy 481 GGDMDNRSELYKYKVKVIEPLGVAPTAKRVRVQREKRAVGLGAMFFGFLGAAGSTWG 540
Db 468 GGNMDNRSELYKYKVKVIEPLGVAPTAKRVRVQREKRAVGLGAMFFGFLGAAGSTWG 527
Qy 541 AASVTLTVQARQLLSGIVQOQSNLLRAEAQOHLQLTVMGKIQOARILAVERYLKDOQ 600
Db 528 AASMTLTVQARQLLSGIVQOQSNLLRAEAQOHLQLTVMGKIQOARILAVERYLKDOQ 587
Qy 601 LLAGWCSGKLICTTTPVNASWSKNKLTLEATWNNMTWMDKEIDNTSLIYSLIESP 660
Db 588 LLAGWCSGKLICTTTPVNASWS-NKSLDNIWNNTWMDREINNTSLIYSLIESQ 646
Qy 661 IQOEKNEBELLELDKAWNLNNWFDISNLLWYIKIFIMIVGGLVGLRIVFVLSIVNRVQ 720
Db 647 NOOEKNEBELLELDKAWNLNNWFDITKWLWYIKIFIMIVGGLVGLRIVFVLSIVNRVQ 706
Qy 721 GYSPISFQRLPAPRDPREPBEIEGGDRDRSGLLVGDLTLIWDLSLCLFSYHR 780
Db 707 GYSPISFQRLPAPRDPREPBEIEGGDRDRSGLLVGDLTLIWDLSLCLFSYHR 766
Qy 781 LRDLILLIVTRIVELLGRGWELTKYWNLLQYWSQELKNSAVLSFNATAIAVAGTDRVI 840
Db 767 LRDLILLIVTRIVELLGRGWELTKYWNLLQYWSQELKNSAVLSFNATAIAVAGTDRVI 826
Qy 841 QVLQVRGRALLHIPTRIQGLERALL 866
Db 827 EVLQRAYRALLHIPTRIQGLERALL 852

RESULT 14
O40222
ID O40222 PRELIMINARY; PRT; 854 AA.
AC O40222;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Env polyprotein.
GN Name:env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=AD8;
RC MEDLINE=96432129; PubMed=8835195;
RA Theodore T.S., England G., Buckler-White A., Buckler C.E.,
  Martin M.A., Peden K.W.;
  "Construction and characterization of a stable full-length macrophage-
  tropic HIV type 1 molecular clone that directs the production of high
  titers of progeny virions.";
```

```
RL AIDS Res. Hum. Retroviruses 12:191-194 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AD8;
RA Theodore T.S., Englund G., Buckler-White A., Buckler C.B.,
RA Martin M.A., Peden K.W.C.;
RA Submitted (May-1997) to the EMBL/GenBank/DBJ databases.
DR AF004394; AAB64170.1; -.
DR HSP; P04578; IDLB.
DR GO:0019031; C:viral envelope; IEA.
DR GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR011010; DNA_brk_join_enz.
DR InterPro: IPR000328; Env_GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 854 AA; 97291 MW; 06C45E69103C6C12 CRC64;

Query Match      84.2%; Score 3886; DB 2; Length 854;
Best Local Similarity 84.3%; Pred. No. 4.5e-279;
Matches 732; Conservative 47; Mismatches 73; Indels 16; Gaps 8;

QY 1 MRVKGIRRYQHWWGCTMLLGLLMICSAATEKLVVTVYGVVWKEATTTLCASDAKAY 60
DB 1 MKVKGIRKYNQHLWKWIMLLGLMLICSAVENLWTVYGVVWKEATTTLCASDAKAY 60
QY 61 DTEAHNVWATHACVPTDPNPQVEVLNVTFNFWKNNVQMHEDIISLWDSQSLKPCVK 120
DB 61 DTEAHNVWATHACVPTDPNPQVEVLNVTFNFWKNNVQMHEDIISLWDSQSLKPCVK 120
QY 121 LTPCLVTLNCTDLRNTNTNNNSSEGTIKGEMKNCSEFNATSTIGDKMQKEYAL 180
DB 121 LTPCLVTLNCTDLRNTNNNS-----SEG--MRGEIKNCSEFNATSTIGDKMQKEYAL 171
QY 181 LYKLDIEPIDNDNTSVRLISCVTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG 240
DB 172 FYRLDVVPIDNDNTSVRLINCNTSTITQACPKVSFEPPIHYCTPAGFAILKCKCKKFG 231
QY 241 KGSCKNVSTVQCTHGRPVVSTOLLNGSLAEVEVIRSENFNTNNAKTIIVQLREPKIN 300
DB 232 TGFCCKNVSTVQCTHGRPVVSTOLLNGSLAEVEVIRSENFNTNNAKTIIVQLKESVEIN 291
QY 301 CSRPNNTTKRSIPMGPGRAFYTTGQIIGDIROAHCNISKNTNNAKQVVEKLGEOFNKT 360
DB 292 CTRPNNTTKRSIHIGPGRAFYTTGQIIGDIROAHCNISRKNNTNNAKQVVEKLGEOFNKN 351
QY 361 K-IVFTNSSGGDPEIVTHSFNCAGEFFYCNNTQTFDSIWSNENGTWNTIRGLNNTGRNDT 419
DB 352 KTVFVNSQSGDPEIVHNSFNCAGEFFYCNSTQTFNSNFWNF-NGTWNLTS-NGTEGNDT 409
QY 420 ITLPCRITKQIINRWQEVGKAYAPPKIGNISCSNNTIGLLTRDGGKDDNSRDGNETFP 479
DB 410 ITLPCRITKQIINRWQEVGKAYAPPKIGNISCSNNTIGLLTRDGG--NNHNDTETFRP 467
QY 480 GGGDMRDNRSELYKYKVVKIEPLGVAPTAKRRVQREERAVG-LGAMFFGLGAAGST 538
DB 468 GGGDMRDNRSELYKYKVVKIEPLGVAPTAKRRVQREERAVGTTGAMFLGLGAAGST 527
QY 539 MGAASVTLTVQAROLSGIVQOQSNLLRAEAQOHLQLTWGIGKQARILAVERYLKD 598
DB 528 MGAASVTLTVQAROLSGIVQOQSNLLRAEAQOHLQLTWGIGKQARVLAVERYLDR 587
QY 599 QQLLGLWGCSGLKICTTTPWPNASWSKNKTLAIWNNMTWMDKEIDNVTSLIYSIER 658
DB 588 QQLLGLWGCSGLKICTTAVPNASWS-NKTLDMWNNMTWMDKEIDNVTGLIYTLIEE 646
QY 659 SPIQEQNEQELLELDKAWNLWNFDSIWNLTWIKIFIMVGLVGLIRIVFVVLSTVNRV 718
DB 647 SQNQEQNEQELLELDKAWNLWNFDSIWNLTWIKIFIMVGLVGLIRIVFVVLSTVNRV 706
QY 719 RQGYSPLSFOTRLPAPRPPDRPEIEEGGDRDRSGLLVGDFLTLIWDLSLCLFSY 778
DB 707 RQGYSPLSFOTRLPAPRPPDRPEIEEGGDRDRSGLLVGDFLTLIWDLSLCLFSY 766
QY 779 HRLRDLILLIVTRIVELLGRRGWEILKYWNLLQYWSQELKNSAVSLFNATAIAVAGTDR 838
DB 767 HRLRDLILLIVARIVELLGRRGWEILKYWNLLQYWSQELKNSAVSLFNATAIAVAGTDR 826
QY 839 VIQVQVGRGALLHIPTIRIQGLERALL 866
DB 827 VIBIVQIRAILHIPTIRIQGLERALL 854

RESULT 15
Q03811 PRELIMINARY; PRT; 853 AA.
ID AC Q03811;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Surface envelope glycoprotein.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91195299; PubMed=2014229;
RA Westervelt P., Gendelman H.E., Ratner L.;
RT "Identification of a determinant within the human immunodeficiency
RT virus 1 surface envelope glycoprotein critical for productive
RT infection of primary monocytes";
RL Proc. Natl. Acad. Sci. U.S.A. 88:3097-3101(1991).
DR EMBL; M60472; AAA45065.1; -.
DR HSP; P20871; 1CE4.
DR GO:0016021; C:integral to membrane; IEA.
DR GO:0019028; C:viral capsid; IEA.
DR GO:0019031; C:viral envelope; IEA.
DR GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR000328; Env_GP41.
DR Pfam: PF00516; GP120; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;
KW Transmembrane.
SQ SEQUENCE 853 AA; 96998 MW; AD2AF21E2B06AD78 CRC64;

Query Match      83.8%; Score 3870.5; DB 2; Length 853;
Best Local Similarity 83.9%; Pred. No. 6.4e-278;
Matches 728; Conservative 52; Mismatches 71; Indels 17; Gaps 8;

QY 1 MRVKGIRRYQHWWGCTMLLGLLMICSAATEKLVVTVYGVVWKEATTTLCASDAKAY 60
DB 1 MKVKGIRKYNQHLWKWIMLLGLMLICSAVENLWTVYGVVWKEATTTLCASDAKAY 60
QY 61 DTEAHNVWATHACVPTDPNPQVEVLNVTFNFWKNNVQMHEDIISLWDSQSLKPCVK 120
DB 61 DTEAHNVWATHACVPTDPNPQVEVLNVTFNFWKNNVQMHEDIISLWDSQSLKPCVK 120
QY 121 LTPCLVTLNCTDLRNTNTNNNSSEGTIKGEMKNCSEFNATSTIGDKMQKEYAL 180
DB 121 LTPCLVTLNCTDLRNTNNNS-----SEG--MRGEIKNCSEFNATSTIGDKMQKEYAL 171
QY 181 LYKLDIEPIDNDNTSVRLISCVTSVITQACPKISFEPPIHYCAPAGFAILKCNCKKFSG 240
DB 172 FYRLDVVPIDNDNTSVRLINCNTSTITQACPKVSFEPPIHYCTPAGFAILKCKCKKFG 231
QY 241 KGSCKNVSTVQCTHGRPVVSTOLLNGSLAEVEVIRSENFNTNNAKTIIVQLREPKIN 300
DB 232 TGFCCKNVSTVQCTHGRPVVSTOLLNGSLAEVEVIRSENFNTNNAKTIIVQLKESVEIN 291
QY 301 CSRPNNTTKRSIPMGPGRAFYTTGQIIGDIROAHCNISKNTNNAKQVVEKLGEOFNKT 360
DB 292 CTRPNNTTKRSIHIGPGRAFYTTGQIIGDIROAHCNISRKNNTNNAKQVVEKLGEOFNKN 351
QY 361 K-IVFTNSSGGDPEIVTHSFNCAGEFFYCNNTQTFDSIWSNENGTWNTIRGLNNTGRNDT 419
```

```

352 KTIVNQSGGDEIIVHSPNCGEFFYCNSTQLFNSWNF-NGTWNLTQS-NGTEGNDT 409
QY
420 ITLPCRIRKQIINWQEVGKAMYAPPIKGNISCSNITGLLLTRDGGKDDNSRDGNETFRP 479
Db
410 ITLPCRIRKQIINWQEVGKAMYAPPIRGQIRCSNITGLILTRDGGTNS--GSEIFRP 466
QY
480 GGGDMRDNRSELYKYKVKIIEPLGVAPTAKRRVVRQREERAVG-LGAMFFGFLGAAGST 538
Db
467 GGGDMRDNRSELYKYKVKIIEPLGVAPTAKRRVVRQREKRAVGTIGAMFLGFLGAAGST 526
QY
539 MGAASVTLTVQAROLLSGIVQOOSNLLRAIEAQOHLQLTVWGIKOLQARILAVERYLKD 598
Db
527 MGAASITLTQARLLLSGIVQOQNLLRAIEAQOHLQLTVWGIKOLQARVLALERYLDR 586
QY
599 QOLLGIWGCCKLICITTVPPNWSKNTLEAIWNNMTWOWDKEDNYTSLIYSLIEE 658
Db
587 QOLLGIWGCCKLICITTVPPNWSKNTLEAIWNNMTWOWDKEDNYTSLIYSLIEE 645
QY
659 SPIQEKNEQELLELDKWNLMNWFDISNWLWYIKIFIMIVGGLVGLRIVFVVLISVNRV 718
Db
646 SQNQEKNEQDILLALDKWASLWNWFDISNWLWYIKIFIMIVGGLIGLRIVFTVLSVNRV 705
QY
719 RQGYSPLSFOTRPLPAPRGDPRPEIEEGGDRDRDRSGLLVGDFLTIWVDLRSCLFSY 778
Db
706 RQGYSPLSFOTRPLPAPRGDPRPEIEEGGDRDRDRSGLLVGDFLTIWVDLRSCLFSY 765
QY
779 HRLRDLIIIVTRIVELLGREGWELKYWNLLQYWSQELKNSAVSLFNATAIYAVAGTDR 838
Db
766 HRLRDLIIIVARIVELLGREGWELKYWNLLQYWSQELKNSAVSLFNATAIYAVAGTDR 825
QY
839 VIOVLQVRGALLHIPTRIPRQGLERALL 866
Db
826 VIEVQRIYRAILHIPTRIPRQGLERULL 853
```

Search completed: April 11, 2005, 11:14:03
Job time : 282.268 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2005, 10:37:31 ; Search time 11.4674 Seconds
(without alignments)
1180.438 Million cell updates/sec

Title: US-09-762-261C-3
Perfect score: 195
Sequence: 1 CSRPNNTKSPMGPRAPFTYTGQIIGDIROAHC 35

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	195	100.0	35	3 AAY79023	Aay79023 HIV-1 RV
2	195	100.0	866	3 AAY79020	Aay79020 HIV-1 env
3	185	94.9	35	2 AAR20883	Aar20883 PND EE378
4	185	94.9	35	2 AAR20953	Aar20953 PND EE669
5	185	94.9	35	2 AAR20853	Aar20853 PND EE322
6	183	93.8	35	2 AAR21134	Aar21134 PND EE228
7	183	93.8	35	2 AAR20911	Aar20911 PND EE533
8	183	93.8	35	2 AAR21135	Aar21135 PND EE228
9	183	93.8	35	2 AAR20912	Aar20912 PND EE533
10	183	93.8	35	2 AAR20910	Aar20910 PND EE533
11	183	93.8	35	2 AAR21136	Aar21136 PND EE228
12	183	93.8	35	2 AAR29113	Aar29113 Group II
13	181	92.8	842	3 AAB69350	Aab69350 HIV-1 non
14	180	92.3	35	2 AAR20945	Aar20945 PND EE665
15	180	92.3	35	2 AAR20947	Aar20947 PND EE665
16	180	92.3	35	2 AAR20946	Aar20946 PND EE665
17	180	92.3	35	2 AAR29110	Aar29110 Group II
18	180	92.3	35	2 AAR29117	Aar29117 Group II
19	179	91.8	35	3 AAB10703	Aab10703 HIV-1 iso
20	179	91.8	35	5 ABG73683	Abg73683 HIV-1 V3-
21	178	91.3	35	2 AAR20936	Aar20936 PND EE639
22	177	90.8	35	8 ADR50777	Adr50777 Antigenic
23	177	90.8	45	8 ADR58150	Adr58150 Novel ant
24	176	90.3	35	2 AAR20881	Aar20881 PND EE378
25	176	90.3	35	2 AAR21092	Aar21092 PND EE37

26	176	90.3	35	2 AAR20837	Aar20837 PND EE304
27	176	90.3	35	2 AAR20840	Aar20840 PND EE308
28	176	90.3	35	2 AAR20851	Aar20851 PND EE320
29	176	90.3	35	2 AAR20882	Aar20882 PND EE378
30	176	90.3	35	2 AAR20941	Aar20941 PND EE661
31	176	90.3	35	2 AAR68030	Aar68030 HIV-1 sub
32	176	90.3	35	2 AAR66751	Aar66751 HIV-1 gpl
33	176	90.3	35	2 AAW07390	Aaw07390 HIV-1 gpl
34	176	90.3	35	2 AAW03356	Aaw03356 Human imm
35	176	90.3	35	2 AAW39914	Aaw39914 V3 sequen
36	176	90.3	35	2 AAY31161	Aay31161 HIV-1 gpl
37	176	90.3	35	3 AAB38961	Aab38961 HIV-1 pep
38	176	90.3	35	4 AAB71933	Aab71933 HIV-1 gpl
39	176	90.3	35	4 AAB6184	Aab6184 HIV-1 V3
40	176	90.3	35	8 ADI26505	Adi26505 Peptide o
41	176	90.3	35	8 ADR50753	Adr50753 Antigenic
42	176	90.3	37	2 AAR41065	Aar41065 HIV-1 com
43	176	90.3	39	2 AAW84056	Aaw84056 HIV-1 Ba-
44	176	90.3	43	2 AAW63735	Aaw63735 HIV-1 JR-
45	176	90.3	45	8 ADR58145	Adr58145 Novel ant

ALIGNMENTS

RESULT 1
AAY79023
ID AAY79023 standard; protein; 35 AA.
XX
AC AAY79023;
XX
DT 12-SEP-2003 (revised)
DT 05-JUN-2000 (first entry)
XX
DE HIV-1 RV strain envelope protein V3 region peptide sequence #2.
XX
KW HIV-1; envelope protein; env; cross reactive anti-serum; antibody;
KW vaccine; antibody; V3 region; detect.
XX
OS Human immunodeficiency virus 1.
XX
PN WO200007631-A1.
XX
PD 17-FEB-2000.
XX
PF 04-AUG-1999; 99WO-US017596.
XX
PR 04-AUG-1998; 98US-0095267P.
XX
(JACK-) JACKSON FOUND HENRY M.
Quinnan GV, Zhang PF;
WPI; 2000-205578/18.
XX
Human immunodeficiency virus-1 envelope protein, useful as a vaccine or immunogenic preparation against HIV-1 infection induces the production of broadly cross-reactive neutralizing anti-serum.
XX
Example 5; Page 27; 54pp; English.
XX
This sequence represents a human immunodeficiency virus type 1 (HIV-1) R2 strain envelope (env) protein V3 region peptide sequence. The invention relates to the HIV-1 env protein or its fragments, which when administered to a mammal, induces the production of broadly cross-reactive neutralizing anti-serum against multiple strains of HIV-1. The HIV-1 env protein and its fragments are useful as vaccines, immunogenic compositions or diagnostic reagents. They can also be used to detect anti-HIV-1/HIV-1 neutralizing antibodies in a sample. The amino acid sequence and its fragments or mRNA sequence are useful for generating antibodies in a mammal. In addition, a recombinant delivery vector containing the env amino acid sequence may also be used as a vaccine. (Updated on 12-SEP-2003 to standardise OS field)

```

XX SQ Sequence 35 AA;
Query Match 100.0%; Score 195; DB 3; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.1e-18;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHC 35
| | | | | | | | | | | | | | | | | | | | |
Db 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHC 35

RESULT 2
AAY79020
ID AAY79020 standard; protein; 866 AA.
XX AC AAY79020;
XX DT 12-SEP-2003 (revised)
XX DT 05-JUN-2000 (first entry)
XX DE HIV-1 envelope protein amino acid sequence.
XX KW HIV-1; envelope protein; env; cross reactive anti-serum; antibody;
XX KW vaccine; antibody; detect.
XX OS Human immunodeficiency virus 1.
XX PN W0200007631-A1.
XX PD 17-FEB-2000.
XX PF 04-AUG-1999; 99WO-US017596.
XX PR 04-AUG-1998; 98US-0095267P.
XX PA (JACK-) JACKSON FOUND HENRY M.
XX PI Quinman GV, Zhang PF;
XX DR WPI; 2000-205578/18.
XX PT Human immunodeficiency virus-1 envelope protein, useful as a vaccine or
PT immunogenic preparation against HIV-1 infection induces the production of
PT broadly cross-reactive neutralizing anti-serum.
XX PS Claim 2; Page 33; 54pp; English.
XX CC This sequence represents a human immunodeficiency virus type 1 (HIV-1)
CC envelope (env) protein amino acid sequence. The invention relates to the
CC HIV-1 env protein or its fragments, which when administered to a mammal,
CC induces the production of broadly cross-reactive neutralising anti-serum
CC against multiple strains of HIV-1. The HIV-1 env protein and its
CC fragments are useful as vaccines, immunogenic compositions or diagnostic
CC reagents. They can also be used to detect anti-HIV-1/HIV-1 neutralizing
CC antibodies in a sample. The amino acid sequence and its fragments or mRNA
CC sequence are useful for generating antibodies in a mammal. In addition, a
CC recombinant delivery vector containing the env amino acid sequence may
CC also be used as a vaccine. (Updated on 12-SEP-2003 to standardise OS
XX field)
XX SQ Sequence 866 AA;
Query Match 100.0%; Score 195; DB 3; Length 866;
Best Local Similarity 100.0%; Pred. No. 3.1e-17;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHC 35
| | | | | | | | | | | | | | | | | | | | |
Db 301 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHC 335

RESULT 3

```

```

AAR20883
XX ID AAR20883 standard; protein; 35 AA.
XX AC AAR20883;
XX DT 01-JUN-1992 (first entry)
XX DE PND EE378-3.
XX KW HIV; PND; principal neutralising determinant; Omp; Neisseria;
XX KW outer membrane proteosome; AIDS; vaccine; envelope gene.
XX OS Human immunodeficiency virus.
XX PN EP471407-A.
XX PD 19-FEB-1992.
XX PF 07-AUG-1991; 91EP-00202025.
XX PR 13-AUG-1990; 90US-00566638.
XX PR 13-AUG-1990; 90US-00566654.
XX PR 13-AUG-1990; 90US-00566656.
XX PA (MERI ) MERCK & CO INC.
XX PI Lewis JA, Davide JP, Waterbury JA;
XX DR WPI; 1992-058471/08.
XX DR N-PSDB; AAQ21298.
XX PT New antigenic conjugate of HIV major neutralisation determinant -
XX PT covalently linked to outer membrane proteosome of Neisseria, useful as
XX PT vaccine and in treating and preventing HIV infections, AIDS and ARC.
XX PS Claim 1; Page 175 + 117-118; 177pp; English.
XX CC The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78
XX CC encode PNDs of HIV (envelope protein) which are useful as immunogens for
XX CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic
XX CC conjugates, PND (of HIV) and Omp (of Neisseria) are prepd. separately,
XX CC then linked by cross-linking agents, monogeneric spacers or bigeneric
XX CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids
XX CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,
XX CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or
XX CC post- exposure to prevent or treat HIV infection or disease, and are
XX CC capable of producing an immune response specific for the immunogen. See
XX CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78
XX SQ Sequence 35 AA;
Query Match 94.9%; Score 185; DB 2; Length 35;
Best Local Similarity 91.4%; Pred. No. 2.5e-17;
Matches 32; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAHC 35
| | | | | | | | | | | | | | | | | | | | |
Db 1 CTRPNNTKRSIPGPGAFYTTGQIIGDIRQAHC 35

RESULT 4
AAR20953
XX ID AAR20953 standard; protein; 35 AA.
XX AC AAR20953;
XX DT 01-JUN-1992 (first entry)
XX DE PND EE669-3.
XX KW HIV; PND; principal neutralising determinant; Omp; Neisseria;
XX KW outer membrane proteosome; AIDS; vaccine; envelope gene.
XX

```


OS Human immunodeficiency virus.

XX EP471407-A.

XX 19-FEB-1992.

XX 07-AUG-1991; 91EP-00202025.

XX 13-AUG-1990; 90US-00566638.

XX 13-AUG-1990; 90US-00566654.

XX 13-AUG-1990; 90US-00566656.

XX (MERI) MERCK & CO INC.

XX Lewis JA, Davide JP, Waterbury JA;

XX WPI; 1992-058471/08.

XX N-PSDB; AAQ21370.

XX New antigenic conjugate of HIV major neutralisation determinant - covalently linked to outer membrane proteosome of Neisseria, useful as vaccine and in treating and preventing HIV infections, AIDS and ARC.

XX Claim 1; Page 175 + 168; 177pp; English.

XX The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78 encode PNDs of HIV (envelope protein) which are useful as immunogens for AIDS vaccines, partic. in the form of conjugates. To prepare antigenic conjugates, PND (of HIV) and Omp (of Neisseria) are prep'd. separately, then linked by cross-linking agents, monogeneric spacers or bigeneric spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala, Glu or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or post- exposure to prevent or treat HIV infection or disease, and are capable of producing an immune response specific for the immunogen. See also AAQ21052-61, AAQ21194-311 and AAQ21321-78

XX Sequence 35 AA;

Query Match 94.9%; Score 185; DB 2; Length 35;
Best Local Similarity 91.4%; Pred. No. 2.5e-17;
Matches 32; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAH 35

Db 1 CTRPNNTKRSIPMGPGAFYTTGQIIGDIRQAH 35

RESULT 5

AAR20853

ID AAR20853 standard; protein; 35 AA.

XX AAR20853;

XX 01-JUN-1992 (first entry)

XX PND EE322-2.

XX HIV; PND; principal neutralising determinant; Omp; Neisseria; outer membrane proteosome; AIDS; vaccine; envelope gene.

XX Human immunodeficiency virus.

XX EP471407-A.

XX 19-FEB-1992.

XX 07-AUG-1991; 91EP-00202025.

XX 13-AUG-1990; 90US-00566638.

XX 13-AUG-1990; 90US-00566654.

XX 13-AUG-1990; 90US-00566656.

XX

PA (MERI) MERCK & CO INC.

XX Lewis JA, Davide JP, Waterbury JA;

XX WPI; 1992-058471/08.

XX N-PSDB; AAQ21268.

XX New antigenic conjugate of HIV major neutralisation determinant - covalently linked to outer membrane proteosome of Neisseria, useful as vaccine and in treating and preventing HIV infections, AIDS and ARC.

XX Claim 1; Page 175 + 93-94; 177pp; English.

XX The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78 encode PNDs of HIV (envelope protein) which are useful as immunogens for AIDS vaccines, partic. in the form of conjugates. To prepare antigenic conjugates, PND (of HIV) and Omp (of Neisseria) are prep'd. separately, then linked by cross-linking agents, monogeneric spacers or bigeneric spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala, Glu or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or post- exposure to prevent or treat HIV infection or disease, and are capable of producing an immune response specific for the immunogen. See also AAQ21052-61, AAQ21194-311 and AAQ21321-78

XX Sequence 35 AA;

Query Match 94.9%; Score 185; DB 2; Length 35;
Best Local Similarity 91.4%; Pred. No. 2.5e-17;
Matches 32; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAH 35

Db 1 CTRPNNTKRSIPMGPGAFYTTGQIIGDIRQAH 35

Db 1 CTRPNNTKRSIPMGPGAFYTTGQIIGDIRQAH 35

Db 1 CTRPNNTKRSIPMGPGAFYTTGQIIGDIRQAH 35

RESULT 6

AAR21134

ID AAR21134 standard; protein; 35 AA.

XX AAR21134;

XX 01-JUN-1992 (first entry)

XX PND EE228-1.

XX HIV; PND; principal neutralising determinant; Omp; Neisseria; outer membrane proteosome; AIDS; vaccine; envelope gene.

XX Human immunodeficiency virus.

XX EP471407-A.

XX 19-FEB-1992.

XX 07-AUG-1991; 91EP-00202025.

XX 13-AUG-1990; 90US-00566638.

XX 13-AUG-1990; 90US-00566654.

XX 13-AUG-1990; 90US-00566656.

XX (MERI) MERCK & CO INC.

XX Lewis JA, Davide JP, Waterbury JA;

XX WPI; 1992-058471/08.

XX N-PSDB; AAQ21232.

XX New antigenic conjugate of HIV major neutralisation determinant - covalently linked to outer membrane proteosome of Neisseria, useful as vaccine and in treating and preventing HIV infections, AIDS and ARC.

XX Claim 1; Page 175 + 64; 177pp; English.

XX CC The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78
 CC encode PNDs of HIV (envelope protein) which are useful as immunogens for
 CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic
 CC conjugates, PND (of HIV) and Omp (of Neisseria) are prepd. separately,
 CC then linked by cross-linking agents, monogeneric spacers or bigeneric
 CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids
 CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,
 CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or
 CC post- exposure to prevent or treat HIV infection or disease, and are
 CC capable of producing an immune response specific for the immunogen. See
 CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78

XX SQ Sequence 35 AA;

Query Match 93.8%; Score 183; DB 2; Length 35;
 Best Local Similarity 91.4%; Pred. No. 4.6e-17;
 Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHC 35
 |:|||||:|||||:|||||:|||||:|||||
 Db 1 CTRPNNTKRSIPGPGRAFYTTGDIIGDIRQAHC 35

RESULT 7

AAQ20911

ID AAR20911 standard; protein; 35 AA.

XX AC AAR20911;

XX DT 01-JUN-1992 (first entry)

XX DE PND EE533-2.

XX HTV; PND; principal neutralising determinant; Omp; Neisseria;
 outer membrane proteosome; AIDS; vaccine; envelope gene.

XX OS Human immunodeficiency virus.

XX PN EP471407-A.

XX PD 19-FEB-1992.

XX PF 07-AUG-1991; 91EP-00202025.

XX PR 13-AUG-1990; 90US-00566638.

XX PR 13-AUG-1990; 90US-00566654.

XX PR 13-AUG-1990; 90US-00566656.

XX PA (MERI) MERCK & CO INC.

XX PI Lewis JA, Davide JP, Waterbury JA;

XX WPI; 1992-058471/08.

XX N-PSDB; AAQ21335.

XX New antigenic conjugate of HIV major neutralisation determinant -
 PT covalently linked to outer membrane proteosome of Neisseria, useful as
 PT vaccine and in treating and preventing HIV infections, AIDS and ARC.

XX PS Claim 1; Page 175 + 140; 177pp; English.

XX CC The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78
 CC encode PNDs of HIV (envelope protein) which are useful as immunogens for
 CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic
 CC conjugates, PND (of HIV) and Omp (of Neisseria) are prepd. separately,
 CC then linked by cross-linking agents, monogeneric spacers or bigeneric
 CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids
 CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,
 CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or
 CC post- exposure to prevent or treat HIV infection or disease, and are
 CC capable of producing an immune response specific for the immunogen. See
 CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78

XX SQ Sequence 35 AA;

Query Match 93.8%; Score 183; DB 2; Length 35;
 Best Local Similarity 91.4%; Pred. No. 4.6e-17;
 Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHC 35
 |:|||||:|||||:|||||:|||||:|||||
 Db 1 CTRPNNTKRSIPGPGRAFYTTGDIIGDIRQAHC 35

RESULT 8

AAQ21135

ID AAR21135 standard; protein; 35 AA.

XX AC AAR21135;

XX DT 01-JUN-1992 (first entry)

XX DE PND EE228-2.

XX HIV; PND; principal neutralising determinant; Omp; Neisseria;
 outer membrane proteosome; AIDS; vaccine; envelope gene.

XX OS Human immunodeficiency virus.

XX PN EP471407-A.

XX PD 19-FEB-1992.

XX PF 07-AUG-1991; 91EP-00202025.

XX PR 13-AUG-1990; 90US-00566638.

XX PR 13-AUG-1990; 90US-00566654.

XX PR 13-AUG-1990; 90US-00566656.

XX PA (MERI) MERCK & CO INC.

XX PI Lewis JA, Davide JP, Waterbury JA;

XX WPI; 1992-058471/08.

XX N-PSDB; AAQ21233.

XX New antigenic conjugate of HIV major neutralisation determinant -
 PT covalently linked to outer membrane proteosome of Neisseria, useful as
 PT vaccine and in treating and preventing HIV infections, AIDS and ARC.

XX PS Claim 1; Page 175 + 65; 177pp; English.

XX CC The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78
 CC encode PNDs of HIV (envelope protein) which are useful as immunogens for
 CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic
 CC conjugates, PND (of HIV) and Omp (of Neisseria) are prepd. separately,
 CC then linked by cross-linking agents, monogeneric spacers or bigeneric
 CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids
 CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,
 CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or
 CC post- exposure to prevent or treat HIV infection or disease, and are
 CC capable of producing an immune response specific for the immunogen. See
 CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78

XX SQ Sequence 35 AA;

Query Match 93.8%; Score 183; DB 2; Length 35;
 Best Local Similarity 91.4%; Pred. No. 4.6e-17;
 Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHC 35
 |:|||||:|||||:|||||:|||||:|||||
 Db 1 CTRPNNTKRSIPGPGRAFYTTGDIIGDIRQAHC 35

RESULT 9
AAR20912
ID AAR20912 standard; protein; 35 AA.

XX AC AAR20912;
XX DT 01-JUN-1992 (first entry)
XX DE PND EE533-3.
XX KW HIV; PND; principal neutralising determinant; Omp; Neisseria;
XX KW outer membrane proteosome; AIDS; vaccine; envelope gene.
XX OS Human immunodeficiency virus.

XX EP471407-A.

XX PN 19-FEB-1992.

XX PF 07-AUG-1991; 91EP-00202025.

XX PR 13-AUG-1990; 90US-00566638.

XX PR 13-AUG-1990; 90US-00566654.

XX PR 13-AUG-1990; 90US-00566656.

XX PA (MERI) MERCK & CO INC.

XX PI Lewis JA, Davide JP, Waterbury JA;

XX DR WPI; 1992-058471/08.

XX DR N-PSDB; AAQ21336.

XX PT New antigenic conjugate of HIV major neutralisation determinant -

XX PT covalently linked to outer membrane proteosome of Neisseria, useful as

XX PT vaccine and in treating and preventing HIV infections, AIDS and ARC.

XX PS Claim 1; Page 175 + 141; 177pp; English.

XX CC The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78

XX CC encode PNDs of HIV (envelope protein) which are useful as immunogens for

XX CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic

XX CC conjugates, PND (of HIV) and Omp (of Neisseria) are prepd. separately,

XX CC then linked by cross-linking agents, monogenic spacers or bigeneric

XX CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids

XX CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,

XX CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or

XX CC post- exposure to prevent or treat HIV infection or disease, and are

XX CC capable of producing an immune response specific for the immunogen. See

XX CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78

XX SQ Sequence 35 AA;

XX Query Match 93.8%; Score 183; DB 2; Length 35;

XX Best Local Similarity 91.4%; Pred. No. 4.6e-17;

XX Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGGIIGDIRQAHG 35

Db 1 CTRPNNTKRSIPMGPGRAFYTTGGIIGDIRQAHG 35

RESULT 10
AAR20910
ID AAR20910 standard; protein; 35 AA.

XX AC AAR20910;
XX DT 01-JUN-1992 (first entry)
XX DE PND EE533-1.
XX KW HIV; PND; principal neutralising determinant; Omp; Neisseria;
XX KW outer membrane proteosome; AIDS; vaccine; envelope gene.

XX OS Human immunodeficiency virus.

XX EP471407-A.

XX PN 19-FEB-1992.

XX PF 07-AUG-1991; 91EP-00202025.

XX PR 13-AUG-1990; 90US-00566638.

XX PR 13-AUG-1990; 90US-00566654.

XX PR 13-AUG-1990; 90US-00566656.

XX PA (MERI) MERCK & CO INC.

XX PI Lewis JA, Davide JP, Waterbury JA;

XX DR WPI; 1992-058471/08.

XX DR N-PSDB; AAQ21334.

XX PT New antigenic conjugate of HIV major neutralisation determinant -

XX PT covalently linked to outer membrane proteosome of Neisseria, useful as

XX PT vaccine and in treating and preventing HIV infections, AIDS and ARC.

XX PS Claim 1; Page 175 + 139-140; 177pp; English.

XX CC The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78

XX CC encode PNDs of HIV (envelope protein) which are useful as immunogens for

XX CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic

XX CC conjugates, PND (of HIV) and Omp (of Neisseria) are prepd. separately,

XX CC then linked by cross-linking agents, monogenic spacers or bigeneric

XX CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids

XX CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,

XX CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or

XX CC post- exposure to prevent or treat HIV infection or disease, and are

XX CC capable of producing an immune response specific for the immunogen. See

XX CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78

XX SQ Sequence 35 AA;

XX Query Match 93.8%; Score 183; DB 2; Length 35;

XX Best Local Similarity 91.4%; Pred. No. 4.6e-17;

XX Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGGIIGDIRQAHG 35

Db 1 CTRPNNTKRSIPMGPGRAFYTTGGIIGDIRQAHG 35

RESULT 11
AAR21136
ID AAR21136 standard; protein; 35 AA.

XX AC AAR21136;

XX DT 01-JUN-1992 (first entry)

XX DE PND EE228-3.

XX KW HIV; PND; principal neutralising determinant; Omp; Neisseria;

XX KW outer membrane proteosome; AIDS; vaccine; envelope gene.

XX OS Human immunodeficiency virus.

XX EP471407-A.

XX PN 19-FEB-1992.

XX PF 07-AUG-1991; 91EP-00202025.

XX PR 13-AUG-1990; 90US-00566638.

XX PR 13-AUG-1990; 90US-00566654.

XX PR 13-AUG-1990; 90US-00566656.

XX (MERI) MERCK & CO INC.
 PA Lewis JA., Davide JP, Waterbury JA;
 PI WPI; 1992-058471/08.
 XX N-PSDB; AAQ21234.
 DR
 XX New antigenic conjugate of HIV major neutralisation determinant -
 PT covalently linked to outer membrane proteosome of Neisseria, useful as
 PT vaccine and in treating and preventing HIV infections, AIDS and ARC.
 XX
 PS Claim 1; Page 175 + 66; 177pp; English.
 XX
 CC The sequences represented in AAQ21054-61, AAQ21194-311 and AAQ21321-78
 CC encode PNDs of HIV (envelope protein) which are useful as immunogens for
 CC AIDS vaccines, partic. in the form of conjugates. To prepare antigenic
 CC conjugates, PND (of HIV) and Omp (of Neisseria) are prep'd. separately,
 CC then linked by cross-linking agents, monogeneric spacers or bigeneric
 CC spacers. The conjugate may comprise 1-50 PNDs. A PND has 35 amino acids
 CC or less, but at least 5, it contains Gly-X-Gly, where X is Pro, Leu, Ala,
 CC Gln or Ser. The conjugates as AIDS or HIV vaccines can be used pre- or
 CC post- exposure to prevent or treat HIV infection or disease, and are
 CC capable of producing an immune response specific for the immunogen. See.
 CC also AAQ21052-61, AAQ21194-311 and AAQ21321-78
 XX
 XX Sequence 35 AA;
 SQ
 Query Match 93.8%; Score 183; DB 2; Length 35;
 Best Local Similarity 91.4%; Pred. No. 4.6e-17;
 Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHC 35
 |:|||||:|||||:|||||:|||||:|||||
 Db 1 CTRPNNTKRSIPGPGRAFYTTGDIIGDIRQAHC 35
 |:|||||:|||||:|||||:|||||:|||||
 RESULT 12
 AAR29113
 ID AAR29113 standard; protein; 35 AA.
 XX
 AC AAR29113;
 XX
 XX 25-MAR-2003 (revised)
 DT 02-APR-1993 (first entry)
 XX
 XX Group II HIV gp120 PND peptide 8926C.
 DE
 XX Principal Neutralising determinant; Human immunodeficiency virus;
 KW vaccine; Robson's analytical method; polymerase chain reaction;
 KW Garnier-Osguthorpe-Robson method; GOR method; secondary structure.
 XX
 OS Human immunodeficiency virus.
 XX
 XX Location/Qualifiers
 FH Region 8..14
 FT /note= "classification of PND peptides is determined by
 FT the predicted secondary structure of this region - see
 FT comments"
 FT 15..18
 FT /note= "conserved PND motif"
 FT
 XX
 XX EP516135-A2.
 XX
 XX 02-DEC-1992.
 PD
 XX 29-MAY-1992; 92BP-00109072.
 XX
 XX 31-MAY-1991; 91JP-00129224.
 XX
 XX (KAGA) CHENO SERO THERAPEUTIC RES INS.
 PA
 XX Eda Y, Shiosaki K, Osatomi K, Tokiyoshi S;

XX WPI; 1992-400517/49.
 XX
 PT Principle neutralising determinant peptide(s) of HIV gp120 protein - used
 PT for diagnosing, preventing and treating HIV infection.
 XX
 XX Example 1; Page 10; 26pp; English.
 PS
 XX DNA encoding HIV PND peptides was PCR amplified using genomic DNA from
 CC HIV-infected peripheral blood mononuclear cells as template. The
 CC amplified fragments were fused to beta-galactosidase coding sequence.
 CC E.coli transformants were cultured to produce the fusion protein. The
 CC expressed PND proteins were divided into groups based on their reactivity
 CC with neutralising antibodies and their amino acid sequence. The amino
 CC acid sequence was analysed using Robson's analytical program for protein
 CC secondary structure. Five groups were identified and 90% of all
 CC previously reported PND peptides were included in 3 main groups (i.e.
 CC Groups I, II and III). Group II PND peptides are those which have the
 CC structure XXBBBX on the amino-terminal side of the GPGR motif (B = beta-
 CC strand structure and X = turn or coil structure). Vaccine preparations
 CC comprising representative peptides from each of the 5 groups can be used
 CC to develop vaccines able to recognise any HIV variant. See AAQ31607-
 CC Q31608, AAR28995-R29000 and AAR29110-R29128. (Updated on 25-MAR-2003 to
 CC correct FN field.)
 XX
 XX Sequence 35 AA;
 SQ
 Query Match 93.8%; Score 183; DB 2; Length 35;
 Best Local Similarity 91.4%; Pred. No. 4.6e-17;
 Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAHC 35
 |:|||||:|||||:|||||:|||||:|||||
 Db 1 CTRPNNTKRSIPGPGRAFYTTGDIIGDIRQAHC 35
 |:|||||:|||||:|||||:|||||:|||||
 RESULT 13
 AAB69350
 ID AAB69350 standard; protein; 842 AA.
 XX
 AC AAB69350;
 XX
 XX 12-SEP-2003 (revised)
 DT 20-APR-2001 (first entry)
 XX
 XX HIV-1 non-subtype B clone 93BR029-4 env protein.
 DE
 XX HIV-1; human immunodeficiency virus; non-subtype B; gag; pol; env; vpu;
 KW vif; vpr; tat; rev; nef; vaccine.
 XX
 OS Human immunodeficiency virus 1.
 XX
 PN WO2000026416-A1.
 XX
 XX 11-MAY-2000.
 PD
 XX 25-OCT-1999; 99WO-US024837.
 PF
 XX 02-NOV-1998; 98US-00184418.
 PR
 XX (UABR-) UAB RES FOUND.
 PA
 XX Hahn BH, Shaw GM, Gao F;
 XX
 XX WPI; 2000-365651/31.
 DR
 XX Novel genomic nucleic acids of non-subtype B human immunodeficiency virus
 PT type 1 useful for detecting and treating AIDS comprises a specific
 PT nucleotide sequence.
 XX
 XX Claim 41; Fig 21; 131pp; English.
 PS
 XX The present in invention provides the protein and coding sequences for a

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2005, 10:57:57 ; Search time 2.65208 Seconds
(without alignments)
1269.790 Million cell updates/sec

Title: US-09-762-261C-3
Perfect score: 195
Sequence: 1 CSRPNNTKRSIPMGFGRAFYTTGQIIGDIRQAHC 35
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	180	92.3	35	2	PC2296
2	176	90.3	35	2	PC2295
3	176	90.3	77	2	S35862
4	176	90.3	110	2	A46410
5	176	90.3	847	2	T09448
6	176	90.3	847	2	S13289
7	175	89.7	76	2	S35859
8	175	89.7	77	2	S35820
9	175	89.7	77	2	S35819
10	174	89.2	443	2	C41621
11	173	88.7	77	2	S35834
12	173	88.7	77	2	S35836
13	173	88.7	77	2	S35835
14	172	88.2	77	2	S35848
15	172	88.2	77	2	S35849
16	172	88.2	77	2	S35857
17	171	87.7	35	2	PC2297
18	171	87.7	46	2	G01847
19	171	87.7	77	2	S35823
20	171	87.7	77	2	S35809
21	171	87.7	110	2	B46410
22	171	87.7	110	2	A46410
23	171	87.7	290	2	S25940
24	171	87.7	445	2	A41621
25	170	87.2	77	2	S35806
26	170	87.2	77	2	S35841
27	170	87.2	77	2	S35816
28	170	87.2	77	2	S35804
29	170	87.2	77	2	S35850

30	170	87.2	843	1	H44001	env polyprotein pr
31	169	86.7	71	2	S42918	env polyprotein -
32	169	86.7	77	2	S35821	envelope protein -
33	168	86.2	77	2	S35824	envelope protein -
34	168	86.2	77	2	S35825	envelope protein -
35	168	86.2	110	2	D46410	envelope protein (
36	168	86.2	852	1	VCLJBR	env polyprotein -
37	167	85.6	71	2	S42904	env polyprotein -
38	167	85.6	71	2	S42905	env polyprotein -
39	167	85.6	71	2	S42898	env polyprotein -
40	167	85.6	77	2	S35815	envelope glycoprote
41	166	85.1	506	2	A40218	envelope protein -
42	165	84.6	77	2	S35800	envelope protein -
43	165	84.6	77	2	S35801	envelope protein -
44	165	84.6	299	2	S60529	envelope polyprote
45	165	84.6	855	1	VCLJAJ	env polyprotein pr

ALIGNMENTS

RESULT 1

PC2296
V3 domain peptide P4346 - human immunodeficiency virus type 1 (fragment)
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 21-Feb-1995 #sequence_revision 05-Apr-1995 #text_change 09-Jul-2004
C:Accession: PC2296

R:Sherefa, K.; Soennerborg, A.; Steinbergs, J.; Saellberg, M.

Biochem. Biophys. Res. Commun. 205, 1658-1664, 1994

A:Title: Rapid grouping of HIV-1 infection in subtypes A to E by V3 peptide serotyping ar

A:Reference number: PC2291; MUID:95110306; PMID:7811250

A:Accession: PC2296

A:Molecule type: protein

A:Residues: 1-35 <SHE>

A:Cross-references: UNIPROT:Q70826; UNIPROT:Q70831

C:Superfamily: Type E retrovirus env polyprotein

Query Match 92.3%; Score 180; DB 2; Length 35;

Best Local Similarity 91.4%; Pred. No. 2.2e-18;

Matches 32; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGFGRAFYTTGQIIGDIRQAHC 35

DB 1 CTRPNNTKRSIHMGFGRAFYTTGQIIGDIRQAHC 35

RESULT 2

PC2295
V3 domain peptide P4317 - human immunodeficiency virus type 1 (fragment)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 21-Feb-1995 #sequence_revision 05-Apr-1995 #text_change 09-Jul-2004

C:Accession: PC2295

R:Sherefa, K.; Soennerborg, A.; Steinbergs, J.; Saellberg, M.

Biochem. Biophys. Res. Commun. 205, 1658-1664, 1994

A:Title: Rapid grouping of HIV-1 infection in subtypes A to E by V3 peptide serotyping ar

A:Reference number: PC2291; MUID:95110306; PMID:7811250

A:Accession: PC2295

A:Molecule type: protein

A:Residues: 1-35 <SHE>

A:Cross-references: UNIPROT:Q8JA81; UNIPROT:Q8JA72; UNIPROT:Q8JA67; UNIPROT:Q8JA75; UNIPR

C:Superfamily: Type E retrovirus env polyprotein

Query Match 90.3%; Score 176; DB 2; Length 35;

Best Local Similarity 88.6%; Pred. No. 8e-18;

Matches 31; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGFGRAFYTTGQIIGDIRQAHC 35

DB 1 CTRPNNTKRSIHMGFGRAFYTTGQIIGDIRQAHC 35

RESULT 3

S35862

envelope protein - human immunodeficiency virus type 1 (strain CSF2951) (fragment)
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
C:Accession: S35862; S35860; S35863
R:Chiodi, F.
submitted to the EMBL Data Library, June 1993
A:Reference number: S35800
A:Accession: S35862
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-77 <CHI>
A:Cross-references: UNIPROT:Q77544; UNIPROT:Q77542; UNIPROT:Q77545; EMBL:Z23246; NID:g313629
A:Experimental source: strain CSF2951 clone3; strain CSF2951 clone4
A:Accession: S35860
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-77 <CH2>
A:Cross-references: EMBL:Z23244; NID:g313631; PIDN:CAA80762.1; PID:g313632; EMBL:Z23247;
A:Experimental source: strain CSF2951 clone11; strain CSF2951 clone4
C:Superfamily: type E retrovirus env polyprotein

Query Match 90.3%; Score 176; DB 2; Length 77;
Best Local Similarity 88.6%; Pred. No. 1.8e-17;
Matches 31; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIIPMGPGRAFYTTGQIIGDIRQAH 35
DB 26 CTRPNNTKSIINIGPGRAFYTTGQIIGDIRQAH 60

RESULT 4
A46410
envelope protein (V3 domain) - human immunodeficiency virus type 1 (fragment)
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A46410
R:McNearney, T.; Hornickova, Z.; Markham, R.; Birdwell, A.; Arens, M.; Saah, A.; Ratner, Proc. Natl. Acad. Sci. U.S.A. 89, 10247-10251, 1992
A:Title: Relationship of human immunodeficiency virus type 1 sequence heterogeneity to A
A:Reference number: A46410; MUID:93066216; PMID:1438212
A:Accession: A46410
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-110 <MCN>
A:Cross-references: UNIPROT:Q9PY06
A:Experimental source: subject S1
A>Note: sequence extracted from NCBI backbone (NCBIP:117724)
C:Superfamily: type E retrovirus env polyprotein

Query Match 90.3%; Score 176; DB 2; Length 110;
Best Local Similarity 88.6%; Pred. No. 2.7e-17;
Matches 31; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIIPMGPGRAFYTTGQIIGDIRQAH 35
DB 37 CTRPNNTKSIINIGPGRAFYTTGQIIGDIRQAH 71

RESULT 5
T09448
envelope glycoprotein - human immunodeficiency virus type 1 (strain JRFL)
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T09448
R:Pang, S.; Vinters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namazie, submitted to the EMBL Data Library, July 1996
A:Reference number: Z16673
A:Accession: T09448
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-847 <PAN>
A:Cross-references: UNIPROT:Q75760; EMBL:U63632; NID:g1465777; PID:g1465781
C:Genetics:

A:Gene: env
C:Superfamily: type E retrovirus env polyprotein

Query Match 90.3%; Score 176; DB 2; Length 847;
Best Local Similarity 88.6%; Pred. No. 2.2e-16;
Matches 31; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIIPMGPGRAFYTTGQIIGDIRQAH 35
DB 293 CTRPNNTKSIINIGPGRAFYTTGQIIGDIRQAH 327

RESULT 6
S13289
env protein - human immunodeficiency virus type 1, HIV-1
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S13289
R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A. Nature 348, 69-73, 1990
A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120 c
A:Reference number: S13288; MUID:91043044; PMID:2172833
A:Accession: S13289
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-847 <OBR>
A:Cross-references: UNIPROT:Q75760
C:Superfamily: type E retrovirus env polyprotein

Query Match 90.3%; Score 176; DB 2; Length 847;
Best Local Similarity 88.6%; Pred. No. 2.2e-16;
Matches 31; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIIPMGPGRAFYTTGQIIGDIRQAH 35
DB 293 CTRPNNTKSIINIGPGRAFYTTGQIIGDIRQAH 327

RESULT 7
S35859
envelope protein - human immunodeficiency virus type 1 (fragment)
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
C:Accession: S35859; S35871
R:Chiodi, F.
submitted to the EMBL Data Library, June 1993
A:Reference number: S35800
A:Accession: S35859
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-76 <CHI>
A:Cross-references: UNIPROT:Q77541; EMBL:Z23243; NID:g313629; PIDN:CAA80761.1; PID:g31363
C:Superfamily: type E retrovirus env polyprotein

Query Match 89.7%; Score 175; DB 2; Length 76;
Best Local Similarity 88.6%; Pred. No. 2.5e-17;
Matches 31; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIIPMGPGRAFYTTGQIIGDIRQAH 35
DB 25 CTRPNNTKSIINIGPGRAFYTTGQIIGDIRQAH 59

RESULT 8
S35820
envelope protein - human immunodeficiency virus type 1 (fragment)
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
C:Accession: S35820; S35822
R:Chiodi, F.
submitted to the EMBL Data Library, June 1993
A:Reference number: S35800
A:Accession: S35820


```

Db      43 CTRPNNNTRKSIISIGPGRAFYTGTGGIIGDIRQAHC 77

RESULT 11
S35834
envelope protein - human immunodeficiency virus type 1 (fragment)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
C;Accession: S35834; S35837
R;Chioldi, F.
submitted to the EMBL Data Library, June 1993
A;Reference number: S35800
A;Accession: S35834
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-77 <CHI>
A;Cross-references: UNIPROT:Q77519; UNIPROT:Q77516; EMBL:Z23216; EMBL:Z23219
C;Superfamily: type E retrovirus env polyprotein

Query Match      88.7%; Score 173; DB 2; Length 77;
Best Local Similarity 88.6%; Pred. No. 4.8e-17;
Matches 31; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 CTRPNNNTRKSIIMGPGRAFYTGTGGIIGDIRQAHC 35
         |:||||| ||||| ||||| ||||| ||||| |||||
Db      26 CTRPNNNTRKSIHMGPGRAPYATGDIIGDIRQAHC 60

RESULT 12
S35836
envelope protein - human immunodeficiency virus type 1 (fragment)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
C;Accession: S35836
R;Chioldi, F.
submitted to the EMBL Data Library, June 1993
A;Reference number: S35800
A;Accession: S35836
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-77 <CHI>
A;Cross-references: UNIPROT:Q77518; EMBL:Z23218
C;Superfamily: type E retrovirus env polyprotein

Query Match      88.7%; Score 173; DB 2; Length 77;
Best Local Similarity 88.6%; Pred. No. 4.8e-17;
Matches 31; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 CTRPNNNTRKSIIMGPGRAFYTGTGGIIGDIRQAHC 35
         |:||||| ||||| ||||| ||||| ||||| |||||
Db      26 CTRPNNNTRKSIHMGPGRAPYATGDIIGDIRQAHC 60

RESULT 13
S35835
envelope protein - human immunodeficiency virus type 1 (fragment)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
C;Accession: S35835
R;Chioldi, F.
submitted to the EMBL Data Library, June 1993
A;Reference number: S35800
A;Accession: S35835
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-77 <CHI>
A;Cross-references: UNIPROT:Q77517; EMBL:Z23217
C;Superfamily: type E retrovirus env polyprotein

Query Match      88.7%; Score 173; DB 2; Length 77;
Best Local Similarity 88.6%; Pred. No. 4.8e-17;
Matches 31; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 CSRPNNTKSIIPMGPGRAFYTGTGIIIGDIRQAH 35
|.|||||
Db 26 CTRPNNTKSIHMGPGRAFATGDIIGDIRQAH 60
|.|||||

RESULT 14

S35848

envelope protein - human immunodeficiency virus type 1 (fragment)

C;Species: human immunodeficiency virus type 1, HIV-1

C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004

C;Accession: S35848

R;Chiodi, F.

submitted to the EMBL Data Library, June 1993

A;Reference number: S35800

A;Accession: S35848

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-77 <CHI>

A;Cross-references: UNIPROT:Q77530; EMBL:Z23232; NID:g313607; PIDN:CAA80750.1; PID:g3136

C;Superfamily: type E retrovirus env polyprotein

Query Match

Best Local Similarity 88.2%; Score 172; DB 2; Length 77;

Matches 30; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIIPMGPGRAFYTGTGIIIGDIRQAH 35
|.|||||Db 26 CTRPNNTKSIHMGPGRAFATGDIIGDIRQAH 60
|.|||||

RESULT 15

S35849

envelope protein - human immunodeficiency virus type 1 (fragment)

C;Species: human immunodeficiency virus type 1, HIV-1

C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004

C;Accession: S35849

R;Chiodi, F.

submitted to the EMBL Data Library, June 1993

A;Reference number: S35800

A;Accession: S35849

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-77 <CHI>

A;Cross-references: UNIPROT:Q77531; EMBL:Z23233; NID:g313609; PIDN:CAA80751.1; PID:g3136

C;Superfamily: type E retrovirus env polyprotein

Query Match

Best Local Similarity 88.2%; Score 172; DB 2; Length 77;

Matches 30; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSRPNNTKSIIPMGPGRAFYTGTGIIIGDIRQAH 35
|.|||||Db 26 CTRPNNTKSIHMGPGRAFATGDIIGDIRQAH 60
|.|||||

Search completed: April 11, 2005, 11:15:28

Job time : 2.65208 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2005, 10:39:11 ; Search time 11.206 Seconds
(without alignments)
1599.394 Million cell updates/sec

Title: US-09-762-261C-3
Perfect score: 195
Sequence: 1 CSRPNNTKSPMGPGRAFYTTGQIIGDIRQAH C 35

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	195	100.0	866	2 Q9WPZ4	Q9wpz4 human immun
2	192	98.5	35	2 Q78198	Q78198 human immun
3	192	98.5	91	2 Q39362	Q39362 human immun
4	192	98.5	91	2 Q69692	Q69692 human immun
5	189	96.9	85	2 Q09556	Q09556 human immun
6	189	96.9	88	2 Q9DXK1	Q9dxk1 human immun
7	189	96.9	171	2 Q6UAS9	Q6uas9 human immun
8	188	96.4	35	2 Q91430	Q91430 human immun
9	188	96.4	35	2 Q76PL9	Q76pl9 human immun
10	188	96.4	35	2 Q76PM0	Q76pm0 human immun
11	188	96.4	35	2 Q76PM1	Q76pm1 human immun
12	188	96.4	35	2 Q76PM2	Q76pm2 human immun
13	188	96.4	35	2 Q76PM3	Q76pm3 human immun
14	188	96.4	35	2 Q76PM4	Q76pm4 human immun
15	188	96.4	35	2 Q76PM5	Q76pm5 human immun
16	188	96.4	35	2 Q76PM6	Q76pm6 human immun
17	188	96.4	35	2 Q76PM7	Q76pm7 human immun
18	188	96.4	35	2 Q76PM8	Q76pm8 human immun
19	188	96.4	35	2 Q76PM9	Q76pm9 human immun
20	188	96.4	35	2 Q76PN1	Q76pn1 human immun
21	188	96.4	35	2 Q76PN2	Q76pn2 human immun
22	188	96.4	35	2 Q76PN3	Q76pn3 human immun
23	188	96.4	35	2 Q76PN4	Q76pn4 human immun
24	188	96.4	35	2 Q77931	Q77931 human immun
25	188	96.4	35	2 Q78039	Q78039 human immun
26	188	96.4	35	2 Q787P3	Q787p3 human immun
27	188	96.4	35	2 Q787P4	Q787p4 human immun
28	188	96.4	35	2 Q787P5	Q787p5 human immun
29	188	96.4	35	2 Q787P6	Q787p6 human immun
30	188	96.4	35	2 Q9QC04	Q9qc04 human immun
31	188	96.4	92	2 Q90205	Q90205 human immun

32	188	96.4	92	2 Q90206	Q90206 human immun
33	188	96.4	92	2 Q79067	Q79067 human immun
34	188	96.4	112	2 Q9DPV9	Q9dpv9 human immun
35	188	96.4	112	2 Q9DPW0	Q9dpw0 human immun
36	188	96.4	112	2 Q9DPW1	Q9dpw1 human immun
37	188	96.4	112	2 Q9DPW2	Q9dpw2 human immun
38	188	96.4	112	2 Q9DPW3	Q9dpw3 human immun
39	188	96.4	113	2 Q71020	Q71020 human immun
40	188	96.4	134	2 Q74948	Q74948 human immun
41	188	96.4	229	2 Q40047	Q40047 human immun
42	188	96.4	229	2 Q40049	Q40049 human immun
43	188	96.4	229	2 Q40051	Q40051 human immun
44	188	96.4	230	2 Q40055	Q40055 human immun
45	188	96.4	235	2 Q40050	Q40050 human immun

ALIGNMENTS

RESULT 1
Q9WPZ4 PRELIMINARY; PRT; 866 AA.
AC Q9WPZ4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Envelope protein.
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP MEDLINE=99236722; PubMed=10221533; DOI=10.1089/088922299311086;
RA Quinlan G.V. Jr., Zhang P.F., Fu D.W., Dong M., Alter H.J.;
RT "Expression and characterization of HIV type 1 envelope protein associated with a broadly reactive neutralizing antibody response.";
RL AIDS Res. Hum. Retroviruses 15:561-570(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Quinlan G.V. Jr., Zhang P.F., Fu D.W., Dong M., Alter H.J.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF128126; AAD40637.3; -.
DR HSSP; P04578; 1DLB.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR PFam; PF00517; GP41; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein; Transmembrane.
KW SEQUENCE 866 AA; 98113 MW; 6E76021833F2EACD CRC64;
SQ SEQUENCE 866 AA; 98113 MW; 6E76021833F2EACD CRC64;
Query Match 100.0%; Score 195; DB 2; Length 866;
Best Local Similarity 100.0%; Pred. No. 4.3e-19;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CSRPNNTKSPMGPGRAFYTTGQIIGDIRQAH C 35
Db 301 CSRPNNTKSPMGPGRAFYTTGQIIGDIRQAH C 335

RESULT 2
Q78198 PRELIMINARY; PRT; 35 AA.
AC Q78198;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope protein (fragment).

GN Name=env;
 OS Human immunodeficiency virus 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Epidermis;
 RX MEDLINE=94309197; PubMed=7545939;
 RA Sala M., Zamburo G., Vartanian J., Marconi A., Bertazzoni U.,
 RA Main-Hobson S.;
 RT "Spatial discontinuities in human immunodeficiency virus type 1
 RT quasiespecies derived from epidermal Langerhans cells of a patient with
 RT AIDS and evidence for double infection.";
 RL J. Virol. 68:5280-5283 (1994).
 DR EMBL; Z34376; CAA84146.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120.
 DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
 KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
 FT NON_TER 1 1
 FT SEQUENCE 35 AA; 3875 MW; 9F5A5B2A698C9ABE CRC64;
 SQ SEQUENCE 35 AA; 3875 MW; 9F5A5B2A698C9ABE CRC64;
 Query Match 98.5%; Score 192; DB 2; Length 35;
 Best Local Similarity 97.1%; Pred. No. 3.9e-20;
 Matches 34; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIIGDIRQAH 35
 Db 1 CTRPNNTKRSIPMGPGRAFYTTGQIIIGDIRQAH 35

RESULT 3
 ID O39362 PRELIMINARY; PRT; 91 AA.
 AC O39362;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Envelope glycoprotein (Fragment).
 GN Name=env;
 OS Human immunodeficiency virus 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98242906; PubMed=9583604;
 RA Lukashov V.V., Op de Coul E.L., Coutinho R.A., Goudsmit J.;
 RA "HIV-1 strains specific for Dutch injecting drug users in
 RT heterosexually infected individuals in The Netherlands.";
 RL AIDS 12:635-641 (1998).
 DR EMBL; AF032157; AAC59382.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
 KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
 FT NON_TER 1 1
 FT SEQUENCE 91 AA; 10363 MW; 0D9690FA2767592E CRC64;
 SQ SEQUENCE 91 AA; 10363 MW; 0D9690FA2767592E CRC64;
 Query Match 98.5%; Score 192; DB 2; Length 91;
 Best Local Similarity 97.1%; Pred. No. 9.9e-20;
 Matches 34; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIIGDIRQAH 35

Db 27 CTRPNNTKRSIPMGPGRAFYTTGQIIIGDIRQAH 61

RESULT 4
 ID Q69692 PRELIMINARY; PRT; 91 AA.
 AC Q69692;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Envelope glycoprotein gp120 (Fragment).
 GN Name=env;
 OS Human immunodeficiency virus 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=R923572;
 RX MEDLINE=96392168; PubMed=8798980;
 RA Lukashov V.V., Kuiken C.L., Boer K., Goudsmit J.;
 RA "HIV type 1 subtypes in The Netherlands circulating among women
 RT originating from AIDS-endemic regions.";
 RL AIDS Res. Hum. Retroviruses 12:951-953 (1996).
 DR EMBL; L76886; AAC37946.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
 KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
 FT NON_TER 1 1
 FT SEQUENCE 91 AA; 10333 MW; 17273BE096D7593E CRC64;
 SQ SEQUENCE 91 AA; 10333 MW; 17273BE096D7593E CRC64;
 Query Match 98.5%; Score 192; DB 2; Length 91;
 Best Local Similarity 97.1%; Pred. No. 9.9e-20;
 Matches 34; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIIGDIRQAH 35
 Db 27 CTRPNNTKRSIPMGPGRAFYTTGQIIIGDIRQAH 61

RESULT 5
 ID O90956 PRELIMINARY; PRT; 85 AA.
 AC O90956;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Gp120 (Fragment).
 GN Name=env;
 OS Human immunodeficiency virus 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Heyndrickx L., Janssens W., Coppens S., Vereecken K., Willems B.,
 RA Franssen K., Colebunders R., Vandenbruaens M., Van der Groen G.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ228189; CAA12819.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR AIDS; Coat protein; Glycoprotein; Transmembrane.
 KW AIDS; Coat protein; Glycoprotein; Transmembrane.
 FT NON_TER 1 1
 FT SEQUENCE 85 AA; 9689 MW; 01BB36F634EA0AFB CRC64;
 SQ SEQUENCE 85 AA; 9689 MW; 01BB36F634EA0AFB CRC64;

Query Match 96.9%; Score 189; DB 2; Length 85;
Best Local Similarity 94.3%; Pred. No. 2.5e-19;
Matches 33; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAH 35
|:|||||||||:|||||||||:|||||||||:|||||
29 CTRPNNTKRSIPMGPGAFYTTGQIIGDIRQAH 63

Db

RESULT 6

Q9DXK1 PRELIMINARY; PRT; 88 AA.

AC Q9DXK1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP MEDLINE=21084776; PubMed=11216936;
RX RA Op de Coul E.L.M., Prins M., Cornelissen M., van der Schoot A.,
RA Boufassa F., Brettle R.P., Hernandez-Aguado L., Schiffer V.,
RA McMenamin J., Rezza G., Robertson R., Zangerle R., Goudsmit J.,
RA Coutinho R.A., Lukashov V.V.;
RT "Using phylogenetic analysis to trace HIV-1 migration among western
RT European injecting drug users seroconverting from 1984 to 1997.";
RL AIDS 15:257-266 (2001).
DR EMBL; AF307295; AAG32463.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000777; GP120.
DR PFam; PF00516; GP120; 1.
FT NON_TER 1 1
FT NON_TER 88 88
SQ SEQUENCE 88 AA; 9925 MW; DD523F6D428DB9E4 CRC64;

Query Match 96.9%; Score 189; DB 2; Length 88;
Best Local Similarity 94.3%; Pred. No. 2.6e-19;
Matches 33; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAH 35
|:|||||||||:|||||||||:|||||||||:|||||
27 CTRPNNTKRSIPMGPGAFYTTGQIIGDIRQAH 61

Db

RESULT 7

Q6UA59 PRELIMINARY; PRT; 171 AA.

AC Q6UA59;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang X.-P., Kahn P.D., Cunha C.B., Dosik D.A., Gawel C., Kaplan M.H.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY375253; AA083712.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR000777; GP120.
DR PFam; PF00516; GP120; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 171 171
SQ SEQUENCE 171 AA; 18865 MW; 46C82ED8FBADEB0B CRC64;

Query Match 96.9%; Score 189; DB 2; Length 171;
Best Local Similarity 94.3%; Pred. No. 5.3e-19;
Matches 33; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAH 35
|:|||||||||:|||||||||:|||||||||:|||||
85 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAH 119

Db

RESULT 8

O91430 PRELIMINARY; PRT; 35 AA.

AC O91430;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98090117; PubMed=9430252;
RA Ida S., Gatanaga H., Shioda T., Nagai Y., Kobayashi N., Shimada K.,
RA Kimura S., Iwamoto A., Oka S.;
RT "HIV type 1 V3 variation dynamics in vivo: long-term persistence of
RT non-syncytium-inducing genotypes and transient presence of syncytium-
RT inducing genotypes during the course of progressive AIDS.";
RL AIDS Res. Hum. Retroviruses 13:1597-1609 (1997).
DR EMBL; AB005412; BAA3329.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000777; GP120.
DR InterPro; IPR011056; Pept_S24_S26_C.
DR PFam; PF00516; GP120; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 35 35
SQ SEQUENCE 35 AA; 3827 MW; 9F47EDF1198BF77E CRC64;

Query Match 96.4%; Score 188; DB 2; Length 35;
Best Local Similarity 94.3%; Pred. No. 1.3e-19;
Matches 33; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGAFYTTGQIIGDIRQAH 35
|:|||||||||:|||||||||:|||||||||:|||||
1 CARPNNTKRSIPGPGAFYTTGQIIGDIRQAH 35

Db

RESULT 9

Q76PL9 PRELIMINARY; PRT; 35 AA.

AC Q76PL9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Envelope protein (Fragment).
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.

```

RC TISSUE=Epidermis;
RX MEDLINE=94309197; PubMed=7545939;
RA Sala M., Zambruno G., Vartanian J., Marconi A., Bertazzoni U.,
RA Wain-Hobson S.;
RT "Spatial discontinuities in human immunodeficiency virus type 1
RT quasiespecies derived from epidermal Langerhans cells of a patient with
RT AIDS and evidence for double infection.";
RL J. Virol. 68:5280-5283(1994).
DR EMBL; Z34489; CAA84253.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR000777; GP120.
DR InterPro; IPR011056; Pept_S24_S26_C.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 35 35
SQ SEQUENCE 35 AA; 3857 MW; 9F5A5B2A698BF77E CRC64;

Query Match 96.4%; Score 188; DB 2; Length 35;
Best Local Similarity 94.3%; Pred. No. 1.3e-19;
Matches 33; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 35
Db 1 CTRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 35

RESULT 10
Q76PM0 PRELIMINARY; PRT; 35 AA.
AC Q76PM0;
RC TISSUE=Epidermis;
RX MEDLINE=94309197; PubMed=7545939;
RA Sala M., Zambruno G., Vartanian J., Marconi A., Bertazzoni U.,
RA Wain-Hobson S.;
RT "Spatial discontinuities in human immunodeficiency virus type 1
RT quasiespecies derived from epidermal Langerhans cells of a patient with
RT AIDS and evidence for double infection.";
RL J. Virol. 68:5280-5283(1994).
DR EMBL; Z34487; CAA84253.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR000777; GP120.
DR InterPro; IPR011056; Pept_S24_S26_C.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 35 35
SQ SEQUENCE 35 AA; 3857 MW; 9F5A5B2A698BF77E CRC64;

Query Match 96.4%; Score 188; DB 2; Length 35;
Best Local Similarity 94.3%; Pred. No. 1.3e-19;
Matches 33; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 35
Db 1 CTRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 35

RESULT 11
Q76PM1 PRELIMINARY; PRT; 35 AA.
AC Q76PM1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)

```

```

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
ON NCBI_TaxID=11676;
RX SEQUENCE FROM N.A.
RC TISSUE=Epidermis;
RX MEDLINE=94309197; PubMed=7545939;
RA Sala M., Zambruno G., Vartanian J., Marconi A., Bertazzoni U.,
RA Wain-Hobson S.;
RT "Spatial discontinuities in human immunodeficiency virus type 1
RT quasiespecies derived from epidermal Langerhans cells of a patient with
RT AIDS and evidence for double infection.";
RL J. Virol. 68:5280-5283(1994).
DR EMBL; Z34437; CAA84207.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR000777; GP120.
DR InterPro; IPR011056; Pept_S24_S26_C.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 35 35
SQ SEQUENCE 35 AA; 3857 MW; 9F5A5B2A698BF77E CRC64;

Query Match 96.4%; Score 188; DB 2; Length 35;
Best Local Similarity 94.3%; Pred. No. 1.3e-19;
Matches 33; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 35
Db 1 CTRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 35

RESULT 12
Q76PM2 PRELIMINARY; PRT; 35 AA.
AC Q76PM2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Envelope protein (Fragment).
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
ON NCBI_TaxID=11676;
RX SEQUENCE FROM N.A.
RC TISSUE=Epidermis;
RX MEDLINE=94309197; PubMed=7545939;
RA Sala M., Zambruno G., Vartanian J., Marconi A., Bertazzoni U.,
RA Wain-Hobson S.;
RT "Spatial discontinuities in human immunodeficiency virus type 1
RT quasiespecies derived from epidermal Langerhans cells of a patient with
RT AIDS and evidence for double infection.";
RL J. Virol. 68:5280-5283(1994).
DR EMBL; Z34435; CAA84205.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR000777; GP120.
DR InterPro; IPR011056; Pept_S24_S26_C.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 35 35
SQ SEQUENCE 35 AA; 3857 MW; 9F5A5B2A698BF77E CRC64;

Query Match 96.4%; Score 188; DB 2; Length 35;
Best Local Similarity 94.3%; Pred. No. 1.3e-19;
Matches 33; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSRPNNTKRSIPMGPGRAFYTTGQIIGDIRQAH 35

```


THIS PAGE BLANK (USPTO)